# ALMOST ISOMETRIC ACTIONS, PROPERTY $(T)$, AND LOCAL RIGIDITY

DAVID FISHER AND GREGORY MARGULIS

Abstract. Let $\Gamma$ be a discrete group with property $(T)$ of Kazhdan. We prove that any Riemannian isometric action of $\Gamma$ on a compact manifold $X$ is locally rigid. We also prove a more general foliated version of this result. The foliated result is used in our proof of local rigidity for standard actions of higher rank semisimple Lie groups and their lattices in [FM2].

One definition of property $(T)$ is that a group $\Gamma$ has property $(T)$ if every isometric $\Gamma$ action on a Hilbert space has a fixed point. We prove a variety of strengthenings of this fixed point properties for groups with property $(T)$. Some of these are used in the proofs of our local rigidity theorems.

## 1. **Introduction**

One of the main results of this paper is the following.

**Theorem 1.1.** *Let $\Gamma$ be a discrete group with property $(T)$. Let $X$ be a compact smooth manifold, and let $\rho$ be a smooth action of $\Gamma$ on $X$ by Riemannian isometries. Then the action is $C^{\infty,\infty}$ locally rigid and $C^{k,k-\kappa}$ locally rigid for every $\kappa>0$ for $k{>}1$.*

We recall the definition of local rigidity.

**Definition 1.2.** *Given a locally compact group $\Gamma$ and a $\Gamma$ action $\rho:\Gamma{\times}X{\to}X$ by $C^k$ diffeomorphisms on a manifold $X$, we say that the action is $C^{k,r}$* locally rigid*, where $r{\leq}k$, if any action $\rho'$ by $C^k$ diffeomorphisms, that is sufficiently $C^k$ close to $\rho$ is conjugate to $\rho$ by a small $C^r$ diffeomorphism. We say an action is $C^{\infty,\infty}$* locally rigid *if any action by $C^\infty$ diffeomorphisms which is sufficiently $C^\infty$ close to $\rho$ is conjugate to $\rho$ by a small $C^\infty$ diffeomorphism.*

**Remark:** Throughout this paper, we assume that $X$, the $\rho(\Gamma)$ invariant metric $g$, and therefore the action $\rho$, are much smoother than any

First author partially supported by NSF grants DMS-9902411 and DMS-0226121 and a PSC-CUNY grant. Second author partially supported by NSF grants DMS-9800607 and DMS-0244406. The authors would also like to thank the FIM at ETHZ for hospitality and support.

perturbation we consider. This assumption is in some sense redundant: given a compact $C^k$ manifold $X$ and a $C^{k-1}$ metric $g$ on $X$, one can show that there is a $C^\infty$ structure on $X$ and a $C^\infty$ metric $g'$ on $X$ invariant under $\mathrm{Isom}(X, g)$.

We topologize the space of $C^k$ actions of $\Gamma$ by taking the compact open topology on the space $\mathrm{Hom}(\Gamma, \mathrm{Diff}^k(X))$. The special case of $C^{k,k}$ local rigidity is exactly local rigidity of the homomorphism $\rho : \Gamma \rightarrow \mathrm{Diff}^k(X)$. Since the $C^\infty$ topology is defined as the inverse limit of the $C^k$ topologies, two $C^\infty$ diffeomorphisms are $C^\infty$ close if they are $C^k$ close for some large $k$. Our proof shows explicitly that, for any $\kappa > 0$, a $C^\infty$ perturbation $\rho'$ of $\rho$ which is sufficiently $C^k$ close to $\rho$ is conjugate to $\rho$ by a $C^\infty$ diffeomorphism which is $C^{k-\kappa}$ close to the identity. Many local rigidity results have been proven for actions of higher rank semisimple groups and their lattices. See the introduction to [FM2] for a more detailed historical discussion.

In fact, Theorem 1.1 follows (though with lower regularity) from a more general foliated version, whose somewhat complicated statement we defer to the next section. We also give two self-contained proofs of Theorem 1.1, since many of the ideas are clearer in that special case, and since one proof gives better regularity in that case. Our foliated result is a principal ingredient in our proof of local rigidity for quasi-affine actions of higher rank semisimple Lie groups and their lattices [FM1, FM2].

Some prior results about local rigidity of isometric actions are known. The question was first investigated for lattices $\Gamma$ in groups $G$, where $G$ is a semisimple Lie group with all simple factors of real rank at least 2. In [Z1], Zimmer proved that any ergodic, volume preserving perturbation of an ergodic, isometric actions of such $\Gamma$ preserves a $C^0$ Riemannian metric $g$. In [Z2], he showed that $g$ was actually smooth. In [Z3], Zimmer extended his result to cover all groups with property $(T)$ of Kazhdan, but still required that the perturbation be ergodic and volume preserving and only constructed an invariant metric rather than a conjugacy. In [Be], Benveniste proves $C^{\infty,\infty}$ local rigidity for isometric actions of cocompact lattices in semisimple groups $G$ as above. As a direct generalization of Zimmer's result, we have the following.

**Theorem 1.3.** *Let $\Gamma$ be a discrete group with property $(T)$. Let $(X, g)$ be a compact Riemannian manifold and let $\rho$ be a smooth action of $\Gamma$ on $X$ preserving $g$. For any $\kappa > 0$, any $\Gamma$ action $\rho'$ which is sufficiently $C^{k+1}$ close to $\rho$ preserves a $C^{k-\kappa}$ Riemannian. Furthermore, if $\rho'$ is a $C^\infty$ action $C^\infty$ close to $\rho$, then $\rho'$ preserves a $C^\infty$ metric.*

Though this theorem is a corollary of Theorem 1.1, we give a simple direct proof of the finite regularity case of Theorem 1.3 in section 4.

A locally compact, $\sigma$-compact group $\Gamma$ has *property* $(T)$ if any continuous isometric action of $\Gamma$ on a Hilbert space has a fixed point. In this paper we generalize this standard fixed point property to a wider class of actions. One can view our results as showing that this fixed point property persists for actions which are perturbations of isometric actions. In fact the fixed point property holds quite generally, even for actions which are only partially defined. Note that it is a theorem of Kazhdan that any discrete group with property $(T)$ is finitely generated and any locally compact, $\sigma$-compact group with property $(T)$ is compactly generated [K].

The definition above is not Kazhdan's original definition of property $(T)$, but is equivalent by work of Delorme and Guichardet [De, Gu]. Kazhdan defined a group $\Gamma$ to have property $(T)$ if the trivial representation of $\Gamma$ is isolated in the Fell topology on the unitary dual of $\Gamma$. For detailed introductions to property $(T)$ see [HV] or [M, Chapter III]. A key step in our proofs is to strengthen standard fixed point properties for groups with property $(T)$. For our foliated local rigidity theorems, we also require an effective method for finding fixed points. One corollary of our general method is a simpler proof of Shalom's result that any finitely generated group with property $(T)$ is a quotient of a finitely presented group with property $(T)$ [S]. See also [Zk] for related results. We also prove a similar result for compactly generated groups with property $(T)$, see Theorem 2.4 below.

We now state a special case of our general fixed point property, that suffices for the proof of $C^{k,k-\frac{1}{2}\dim(X)}$ local rigidity.

**Definition 1.4.** *Let* $\varepsilon{\geq}0$ *and* $Z$ *and* $Y$ *be metric spaces. Then a map* $h : Z{\to}Y$ *is an* $\varepsilon$-almost isometry *if*

$$(1-\varepsilon)d_Z(x,y){\leq}d_Y(h(x),h(y)){\leq}(1+\varepsilon)d_Z(x,y)$$

*for all* $x,y{\in}Z$.

The reader should note that an $\varepsilon$-almost isometry is a bilipschitz map. We prefer this notation and vocabulary since it emphasizes the relationship to isometries.

**Definition 1.5.** *Given a group* $\Gamma$ *acting on a metric space* $X$*, a compact subset* $K$ *of* $\Gamma$ *and a point* $x{\in}X$*. The number* $\sup_{k\in K} d(x,k{\cdot}x)$ *is called the* $K$*-displacement of* $x$ *and is denoted* $\mathrm{disp}_K(x)$.

**Theorem 1.6.** *Let* $\Gamma$ *be a locally compact,* $\sigma$*-compact group with property* $(T)$ *and* $K$ *a compact generating set. There exist positive constants*

*$\varepsilon$ and $D$, depending only on $\Gamma$ and $K$, such that for any continuous action of $\Gamma$ on a Hilbert space $\mathcal{H}$ where $K$ acts by $\varepsilon$-almost isometries there is a fixed point $x$; furthermore for any $y$ in $X$, the distance from $y$ to the fixed set is not more than $D\,\mathrm{disp}_K(y)$.*

We note that in most of our applications, the $\varepsilon$-almost isometric action to which we apply Theorem 1.6 and its generalizations are linear, and therefore automatically has a fixed point, the 0 vector. The importance of the final claim in Theorem 1.6 then becomes clear: we have a linear relationship between the distance from a point to the fixed set and the $K$-displacement of the point. In our applications this is used to find non-zero fixed vectors in certain linear actions. That the fixed vector is close to a particular vector with particular prescribed properties is also central to the proof.

Preliminary forms of Theorems 1.1 and 1.6 were announced by the second author at a talk in Jerusalem in 1997.

Theorem 1.6 is proved by contradiction. We assume the existence of a sequence of $\varepsilon$-almost isometric actions not satisfying the conclusion of that theorem, with $\varepsilon$ going to zero. One then constructs a limit action which is isometric and therefore must have fixed points. One then uses a quantitative strengthening of the fixed point property to show that actions "close enough" to the limit action must have fixed points as well. In this article we use ultra-filters and ultra-limits to produce the limit action which considerably simplifies earlier versions of the argument. The argument is further simplified by our use of a stronger quantitative strengthening of the fixed point property which is, in fact, an iterative method for producing fixed points. For $\Gamma$ not discrete, additional difficulties arise from the fact that the limit action is not a priori continuous.

Though the approaches and applications are different, our strengthenings of property $(T)$ are related to the strengthenings discussed by Gromov in [Gr2]. In particular, in section 3.13B, Gromov outlines a proof of Theorem 1.6, though only for a certain class of "random" infinite, discrete groups with property $(T)$ and only for affine $\varepsilon$-almost isometric actions. See Appendix D.2 for further discussion.

Our original approach to proving Theorem 1.1 remains incomplete, though the idea is instructive. Given an isometric action $\rho$ of $\Gamma$ on a compact manifold $X$ and a perturbation $\rho'$ of $\rho$, a conjugacy is a diffeomorphism $f : X \to X$ such that $\rho(\gamma) \circ f = f \circ \rho'(\gamma)$ for all $\gamma$ in $\Gamma$. Rearranging, the conjugacy is a fixed point for the $\Gamma$ action on the group $\mathrm{Diff}^k(X)$ of diffeomorphisms of $X$ defined by $f \to \rho(\gamma) \circ f \circ \rho'(\gamma)^{-1}$. Ideally we would parameterize diffeomorphisms of $X$ locally as a Hilbert

space and then use Theorem 2.3 below, a generalization of Theorem 1.6 for partially defined actions, to find a fixed point or conjugacy. This approach does not work, see Appendix D.1 for further discussion.

Our two proofs of Theorem 1.1 have distinct advantages. We outline here the simpler one which allows us to prove our general foliated result. We discuss here only the result that uses Theorem 1.6 and only indicate a proof of $C^{k,k-\dim(X)-1}$ local rigidity. Even combined with arguments below which improve regularity, this proof requires the loss of $1+\kappa$ derivatives. The precise regularity of Theorem 1.1 requires a different argument and requires stronger assumptions on the action in the foliated case. In subsection 5.1 we include the other proof of Theorem 1.1 but only briefly indicate how, and when, it can be foliated.

Given a compact Riemannian manifold $X$, there is a canonical construction of a Sobolev inner product on $C^k(X)$ such that the Sobolev inner product is invariant under isometries of the Riemannian metric, see section 4 below. We call the completion of $C^k(X)$ with respect to the metric induced by the Sobolev structure $L^{2,k}(X)$. Given an isometric $\Gamma$ action $\rho$ on a manifold $M$ there may be no non-constant $\Gamma$ invariant functions in $L^{2,k}(X)$. However, if we pass to the diagonal $\Gamma$ action on $X{\times}X$, then any function of the distance to the diagonal is $\Gamma$ invariant and, if $C^k$, is in $L^{2,k}(X{\times}X)$.

We choose a smooth function $f$ of the distance to the diagonal in $X{\times}X$ which has a unique global minimum at $x$ on $\{x\}{\times}X$ for each $x$, and such that any function $C^2$ close to $f$ also has a unique minimum on each $\{x\}{\times}X$. This is guaranteed by a condition on the Hessian and the function is obtained from $d(x,y)^2$ by renormalizing and smoothing the function away from the diagonal. This implies $f$ is invariant under the diagonal $\Gamma$ action defined by $\rho$. Let $\rho'$ be another action $C^k$ close to $\rho$. We define a $\Gamma$ action on $X{\times}X$ by acting on the first factor by $\rho$ and on the second factor by $\rho'$. For the resulting action $\bar{\rho}'$ of $\Gamma$ on $L^{2,k}(X{\times}X)$ and every $k{\in}K$, we show that $\bar{\rho}'(k)$ is an $\varepsilon$-almost isometry and that the $K$-displacement of $f$ is a small number $\delta$, where both $\varepsilon$ and $\delta$ can be made arbitrarily small by choosing $\rho'$ close enough to $\rho$. Theorem 1.6 produces a $\bar{\rho}'$ invariant function $f'$ close to $f$ in the $L^{2,k}$ topology. Then $f'$ is $C^{k-\dim(X)}$ close to $f$ by the Sobolev embedding theorems and if $k-\dim(X){\geq}2$, then $f$ has a unique minimum on each fiber $\{x\}{\times}X$ which is close to $(x,x)$. We verify that the set of minima is a $C^{k-\dim(X)-1}$ submanifold and, in fact, the graph of a conjugacy between the $\Gamma$ actions on $X$ defined by $\rho$ and $\rho'$.

In the context of Theorem 1.6, we can prove that given any vector $v$, one can produce any invariant vector $v_0$ by an iterative method of "averaging over balls in $\Gamma$". The proofs of the $C^\infty$ cases of a Theorems 1.1 and 1.3 rely on this iterative method and additional estimates. If our perturbation $\rho'$ is $C^k$ close to $\rho$, using this iterative method, convexity estimates on derivatives and estimates on compositions we produce a sequence of $C^\infty$ diffeomorphisms which converge to conjugacy in the $C^l$ topology for some $l > k$. We then apply an additional iterative argument loosely inspired by the KAM method, to produce the actual $C^\infty$ conjugacy. For a discussion of the relation between our work and the KAM method, see Appendix D.1. The proof of $C^{k,k-\kappa}$ local rigidity for any $\kappa > 0$ and of the lower loss of regularity in Theorem 1.3 follow from a somewhat technical result which allows us to show that the iterative procedure defined by "averaging over balls" also converges in $L^p$ type Sobolev spaces where $p > 2$. We defer statements of these results to subsection 2.2. Once one replaces standard consequences of property $(T)$ with an observation of Bader and Gelander [BFGM], the proof of this result is similar to the proofs of our results concerning actions on Hilbert spaces.

The proof of the foliated generalization of Theorem 1.1 follows a similar outline, but is more difficult at several steps. The choice of initial invariant function is slightly more complicated since leaves of the foliation are generally non-compact. The absence of a natural topology on the set of pairs of points on the same leaf forces us to work on the holonomy groupoid of the foliation. Since we need to work in a Sobolev space defined by only taking derivatives along the leaves of the foliation, having small norm in this topology on functions only gives a good $C^k$ estimate on the conjugacy on a set $S$ of large measure. To guarantee that the orbit of $S$ covers all of $X$ , we use our effective method of producing $\tilde{f}$ from $f$ by "averaging over balls" in $\Gamma$.

**Plan of the paper:** In section 2 we make the necessary definitions and state our general results. First, in subsection 2.1 we discuss various generalizations of Theorem 1.6 for actions on Hilbert spaces. Second in subsection 2.2 we discuss various generalizations of Theorem 1.6 for actions on more general Banach spaces. Then in subsection 2.3 we describe our foliated generalization of Theorem 1.1. Subsection 3.1 and 3.2 contain preliminaries on, respectively, groups with property $(T)$ and limits of actions. We then proceed to prove the results from subsections 2.1 and 2.2 in subsections 3.3 and 3.4 respectively. Section 4 gives an explicit construction of various Sobolev metrics on various spaces of tensors on Riemannian manifolds and more general spaces

with Riemannian foliations. Section 4 also contains a proof of Theorem 1.3. Section 5 contains two proofs of Theorem 1.1. In section 6, we prove the $C^\infty$ case of Theorem 1.1. Section 7 contains some additional background on foliations, a discussion of the holonomy groupoid of a foliation, and a proof of the foliated generalization of Theorem 1.1. On first reading the paper, the reader may wish to skip subsections 2.2, 2.3, 3.4, read 4 assuming $p = 2$ everywhere and assuming that the foliation is by a single leaf and then read subsection 5.1. This allows the reader to read the proof of Theorem 1.1 for the $C^{k,k-\frac{\dim(X)}{2}}$ case, before beginning to study the techniques for improving regularity and/or the, significantly more complicated, formulation and proof of the foliated version.

*Acknowledgements:* The authors would like to thank: Dmitry Dolgopyat, Bassam Fayad and Raphael Krikorian for useful conversations concerning the KAM method and convexity of derivatives; Uri Bader and Tsachik Gelander for sharing their observations on property $(T)$ and $L^p$ spaces and for useful conversations concerning Banach spaces and positive definite functions; and Tim Riley for useful conversations concerning ultrafilters, ultralimits and ultraproducts of metric spaces. We also thank the referees for copious helpful comments that considerably improved the exposition.

## 2. **Definitions and statements of main results**

In this section we give the necessary definitions and state our general results. The first subsection is devoted to general results on actions and partially defined actions of groups with property $(T)$ on Hilbert spaces. The second subsection concerns generalizations of some of these results to more general Banach spaces. The third subsection concerns the foliated version of Theorem 1.1.

**On Constants:** Throughout this paper, we use a convention to simplify the specification of which constants depend on which other choices. When introducing a constant $C$, we will use the notation $C = C(\alpha, \beta, S)$ to specify that $C$ depends on choices of $\alpha, \beta$ and $S$. We make one exception to this rule: as most constants in this paper depend on a choice of a group $\Gamma$ and a generating $K$, we will always leave this dependence implicit. The few cases where constants do not actually depend on an ambient choice of $\Gamma$ and $K$ are clear from context as they appear in statements where $\Gamma$ and $K$ are irrelevant.

2.1. **Fixed points for actions of groups with property $(T)$ on Hilbert spaces.** Throughout this subsection $\Gamma$ will be locally compact, $\sigma$-compact, group generated by a fixed compact subset $K$, which contains a neighborhood of the identity. It follows from work of Kazhdan [K] that any locally compact, $\sigma$-compact $\Gamma$ with property $(T)$ is compactly generated. Given any compact generating set $C$, a simple Baire category argument shows that $C^s$ contains a neighborhood of the identity for some positive integer $s$. (Given a subset $K$ of a group $\Gamma$, we write $K^s$ for the set of all elements of $\Gamma$ that can be written as a product of $s$ elements of $K$.)

Theorem 1.6 suffices to prove $C^{k,k-\frac{\dim(X)}{2}}$ local rigidity in Theorem 1.1. To obtain better finite regularity, a $C^{\infty,\infty}$ local rigidity result, and to prove our more general results, we will need more precise control over how one obtains an invariant vector from an almost invariant vector. As noted above, most of the applications of our results are to the case where the $\varepsilon$ almost isometric actions are actually linear representations. Since the statements of our results do not simplify in any useful way in that setting, we leave it to the interested reader to state the special cases.

Fix a (left) Haar measure $\mu_\Gamma$ on $\Gamma$. We let $\mathcal{U}(\Gamma)$ denote the set of continuous non-negative functions $h$ with compact support on $\Gamma$ with $\int_\Gamma h d\mu_\Gamma = 1$. Given $h \in \mathcal{U}(\Gamma)$ and an action $\rho$ of $\Gamma$ on a Hilbert space $\mathcal{H}$, we can define an operator $\rho(h)$ on $\mathcal{H}$. Let $\rho(h)v = \int_\Gamma \rho(\gamma)(h(\gamma)v)d\mu_\Gamma$. It is straightforward to see that $\int_\Gamma h d\mu_\Gamma = 1$ implies that this definition does not depend on the choice of basepoint in $\mathcal{H}$. When the action $\rho$ is not affine, $\rho(h)$ is not necessarily an affine transformation. We let $\mathcal{U}_2(\Gamma)$ be the subset of functions $h \in \mathcal{U}(\Gamma)$ such that $h > 0$ on $K^2$. We denote by $f * g$ the convolution of integrable functions $f$ and $g$. Note that if $f, g \in \mathcal{U}(\Gamma)$ then so is $f * g$. Given a positive integer $d$, we denote by $f^{*d}$ the $d$-fold convolution of $f$ with itself. More generally, if $\mathcal{P}(\Gamma)$ is the set of probability measures on $\Gamma$ and $\nu \in \mathcal{P}(\Gamma)$, we can also define $\rho(\nu)v = \int_\Gamma \rho(\gamma)v d\nu$. Note that $\mathcal{U}(\Gamma) \subset \mathcal{P}(\Gamma)$ and that this definition generalizes the one above.

We now state a theorem which implies Theorem 1.6. This theorem implies that iterates of certain averaging operators converge to a bounded projection onto the set of fixed points for the action.

**Theorem 2.1.** *If $\Gamma$ has property $(T)$ and $f \in \mathcal{U}_2(\Gamma)$ and $0 < C < 1$, there exists $\varepsilon > 0$, and positive integers $m = m(C, f)$ and $M = M(C, f)$, such that, letting $h = f^{*m}$, for any Hilbert space $\mathcal{H}$, any continuous action of $\Gamma$ on $\mathcal{H}$ such that $K$ acts by $\varepsilon$-almost isometries, and any $x \in \mathcal{H}$ we have*

(1) $d_{\mathcal{H}}(x, \rho(h)(x)) \leq M \operatorname{disp}_K(x)$
(2) $\operatorname{disp}_K(\rho(h)(x)) \leq C \operatorname{disp}_K(x)$.

**Remark:** In this theorem choosing smaller values of $C$ increases the value of $m$. The number $M$ is the least integer with $\operatorname{supp}(h) = \operatorname{supp}(f^{*m}) \subset K^M$.

*Proof of Theorem 1.6 from Theorem 2.1.* The hypotheses of the theorems are almost identical. Since the $\Gamma$ action in Theorem 1.6 is continuous, it follows that every point $x \in \mathcal{H}$ has finite $K$-displacement. Given a point $x \in \mathcal{H}$ with $K$-displacement $\delta$, we look at the sequence $y_n = \rho(h)^n(x)$. Theorem 2.1 implies that the $\operatorname{disp}_K(y_n) \leq C^n \delta$ and that $d(y_n, y_{n+1}) \leq MC^n \delta$. This implies that $y_n$ is a Cauchy sequence and $y = \lim_{n\to\infty} y_n$ clearly has $K$ displacement zero. Letting $D = \sum_{i=1}^{\infty} MC^n = \frac{MC}{1-C}$, then $d_{\mathcal{H}}(x, y) < D\delta$ which completes the proof. $\square$

We now make precise our notion of a partially defined action. By $B(x, r)$ we denote the ball around $x$ of radius $r$.

**Definition 2.2.** *Let $X$ a metric space and fix a point $x \in X$. Given $r, s, \varepsilon, \delta > 0$, we call a map $\rho : K^s \times B(x, r) \to X$ an $(r, s, \varepsilon, \delta, K)$-*almost action of $\Gamma$ on $X$ at $x$ *if the following conditions hold.*

(1) *For each $d \in K^s$, the map $\rho(d, \cdot) : B(x, r) \to X$ is an $\varepsilon$-almost isometry.*
(2) $\operatorname{disp}_K(x) < \delta$.
(3) *With the notation $\rho(d, z) = \rho(d)z$, if $ab, a$ and $b$ are in $K^s$ then $\rho(a)(\rho(b)y) = \rho(ab)y$ whenever $\rho(b)y$ is in $B(x, r)$.*

When $K$ is fixed, we sometimes abbreviate the above notation by calling an $(r, s, \varepsilon, \delta, K)$-almost action an $(r, s, \varepsilon, \delta)$-almost action. We denote by a $(\infty, s, \varepsilon, \delta, K)$-almost action the case when $B(x, r)$ in the definition above can be replaced by $X$. The following theorem now produces fixed points for partially defined actions on Hilbert spaces that are "close enough" to isometric ones.

**Theorem 2.3.** *If $\Gamma$ has property $(T)$ and $\delta_0 > 0$ there exist $\varepsilon > 0, D > 0$, a positive integer $s$, and $r = r(\delta_0) > 0$ such that for any Hilbert space $X$, any $\delta \leq \delta_0$ and any $x \in X$, any continuous $(r, s, \varepsilon, \delta, K)$-action of $\Gamma$ on $X$ at $x$ has a fixed point. Furthermore, the distance from the fixed point to $x$ is not more than $D\delta$.*

Fixing $\delta_0$ is only necessary as a normalization. If we compose a given action with a homothety, we may always assume $\delta_0$ is 1. The constants $s$ and $\varepsilon$ remain unchanged by this process, but $r$ becomes $\delta_0 r$. The utility of considering partially defined actions is illustrated

by our proof of the observation of Shalom stated in the introduction. In fact, we prove the following generalization, which is used in section 7. For background on the notion of a compact presentation see [Ab].

**Theorem 2.4.** *Let* $\Gamma$ *be a locally compact,* $\sigma$*-compact group with property* $(T)$*. Then* $\Gamma$ *is a quotient of a compactly presented locally compact,* $\sigma$*-compact group with property* $(T)$*.*

*Proof.* As remarked above, by work of Kazhdan, $\Gamma$ is compactly generated, and we fix a compact generating set $K$. Possibly after replacing $K$ with a power of $K$, we can assume that $K$ contains a neighborhood of the identity. The group $\Gamma$ is the quotient of the group $\Gamma'$ generated by $K$ satisfying all relations of $\Gamma$ of the form $xy = z$ where $x, y, z \in K^s$. Since $\Gamma'$ satisfies all the relations contained in $K$, we can topologize $\Gamma'$ so that the projection $\Gamma' \to \Gamma$ is a homeomorphism in a neighborhood of the identity and therefore $\Gamma'$ is locally compact and $\sigma$-compact. We believe that this fact is known, but state it as Proposition C.1 in Appendix C where we also sketch a proof, as we did not find a reference in the literature. Since a continuous isometric $\Gamma'$ action is a continuous $(\infty, s, 0, \delta)$-action of $\Gamma$ at $x$ where $\delta$ is the $K$-displacement of $x$, Theorem 2.3 implies that, if we choose $s$ large enough, $\Gamma'$ has property $(T)$. It is clear that $\Gamma'$ is compactly presented. $\square$

**Remarks:**

(1) Theorem 2.4 is used in the proof of Theorem 2.11, the foliated generalization of Theorem 1.1. It is used to show that an action of a locally compact group with property $(T)$ on a compact foliated space lifts to an action on the holonomy groupoid of the foliation.
(2) It is also possible to prove Theorem 2.11 directly from Theorem 2.5 and Corollary 2.8 below.

We now state a generalization of Theorem 2.1 which implies Theorem 2.3. We note that the operator $\rho(h)$ is well defined for a $(r, s, \varepsilon, \delta, K)$-action $\rho$, provided the support of $h$ is contained in $K^s$.

**Theorem 2.5.** *If* $\Gamma$ *has property* $(T)$ *and* $f \in \mathcal{U}_2(\Gamma), 0 < C < 1$ *and* $\delta_0 > 0$ *there exist* $r = r(\delta_0, f, C) > 0$ *and* $\varepsilon > 0$ *and positive integers* $m = m(f, C), s = s(f, C)$ *and* $M = M(f, C)$ *such that, letting* $h = f^{*m}$*, for any Hilbert space* $X$*, any* $\delta \leq \delta_0$*, any* $x \in X$*, and any continuous* $(r, s, \varepsilon, \delta, K)$*-action* $\rho$ *of* $K$ *on* $X$ *at* $x$ *we have*

(1) $d_{\mathcal{H}}(x, \rho(h)(x)) \leq M \operatorname{disp}_K(x)$;
(2) $\operatorname{disp}_K(\rho(h)(x)) \leq C \operatorname{disp}_K(x)$.

*Proof of Theorem 2.3 from Theorem 2.5.* The proof is almost identical to the proof of Theorem 1.6 from Theorem 2.1. One point requires additional care: if $r_0$ and $C$ are the constants given by Theorem 2.5, we need to take $r$ in Theorem 2.3 to be at least $r_0 + M\sum_{i=1}^{\infty} C_1^i$. This insures that we can apply Theorem 2.5 to each $\rho(h)^i(x)$ successively, since it implies that $\rho$ defines an $(r, s, \varepsilon, C^i\delta)$-action on $\mathcal{H}$ at $\rho(h)^i(x)$. $\square$

**Remarks:**

(1) That Theorem 2.5 implies Theorem 2.1 is clear from the definitions. Section 3 is devoted to the proof of Theorem 2.5.

(2) For most of our dynamical applications Theorem 2.1 suffices. However, as remarked above, we need Theorem 2.4, and therefore Theorem 2.3, for the proof of Theorem 2.11. As remarked above, one can also prove Theorem 2.11 using Theorem 2.5 in place of the combination of Theorem 2.1 and Theorem 2.4.

2.2. **Property $(T)$ and uniformly convex Banach spaces.** In this subsection we describe some generalizations of the results in the previous subsection to non-Hilbertian Banach spaces. Throughout this subsection $\Gamma$ and $K$ will be as in the previous subsection. For $1 < p \leq 2$, we will call a Banach space $\mathcal{B}$ a *generalized $L^p$ space*, if the function $\|x\|^p$ is negative definite on $\mathcal{B}$ or equivalently if $\exp(-t\|x\|^p)$ is positive definite for all $t > 0$. A theorem of Bretagnolle, Dacunba-Castelle and Krivine implies that any generalized $L^p$ space is a closed subspace of an $L^p$ space, see [BL, Theorem 8.9]. For $q > 2$, we will called a Banach space $\mathcal{B}$ a *generalized $L^q$ space* if the dual of $\mathcal{B}$ is a generalized $L^p$ space where $\frac{1}{p} + \frac{1}{q} = 1$. Given a finite dimensional Euclidean space $V$ with Euclidean norm $\|\cdot\|_V$ and a measure space $(S, \mu)$ we define a norm on measurable maps $f : S \to V$ by $\|f\|^p = \int_S \|f(s)\|_V^p d\mu$ and let $L^p(S, \mu, V)$ be the space of equivalence classes of maps $f$ with finite norm. If $\dim(V) = n$ and $1 < p < \infty$, we will call $L^p(S, \mu, V)$ a *Banach space of type $L_n^p$*. It is easy to verify that if $\frac{1}{p} + \frac{1}{q} = 1$ then the dual of a Banach space of type $L_n^p$ is a Banach space of type $L_n^q$. It is also easy to verify, for $1 < p < \infty$, that a Banach space of type $L_n^p$ is a generalized $L^p$ space. For $p \leq 2$ this is shown by embedding $L^p(S, \mu, V)$ into $L^p(S \times S_1(V), \mu \times \nu)$ where $S_1(V)$ is the unit sphere in $V$ and $\nu$ is (normalized) Haar measure. For $p > 2$ it is immediate from the definitions.

We now state a variant of Theorem 2.1 for affine actions on Banach spaces.

**Theorem 2.6.** *If $\Gamma$ has property $(T)$ and $f \in \mathcal{U}_2(\Gamma), \delta_0 > 0, 0 < C < 1$, and $\eta > 0$, there exist $\varepsilon = \varepsilon(\eta) > 0$ positive integers $m = m(f, C, \eta), s = s(f, C, \eta)$ and numbers $r = r(\delta_0, \eta, f, C) > 0$ and $M = M(f, C, \eta)$ such that, letting $h = f^{*m}$, for any generalized $L^p$ space $\mathcal{B}$ where $1+\eta < p \leq 2$, any $\delta \leq \delta_0$, any $x \in \mathcal{B}$, and any continuous affine $(r, s, \varepsilon, \delta, K)$-action $\rho$ of $K$ on $\mathcal{B}$ at $x$ we have*

1. $d_{\mathcal{H}}(x, \rho(h)(x)) \leq M \operatorname{disp}_K(x)$;
2. $\operatorname{disp}_K \rho(h)(x) \leq C \operatorname{disp}_K(x)$.

Though they are only concerned with finding fixed points and do not discuss the iterative method, the special case of Theorem 2.6 for (globally defined) isometric actions is essentially contained in [BFGM]. Modulo that fact, the proof of this theorem is quite similar to the proof of Theorem 2.5. In [BFGM], it is also proven that a version of Theorem 2.1 holds for unitary representations in $L^p$ spaces with $2 < p < \infty$ using a simple duality argument. (Once again they only find fixed points, and do not describe the iterative method for finding them.) It would be interesting to know if this is true for representations which are only "almost unitary" and $p > 2$, but we only need a weaker statement for our applications, which we now deduce. We first define the relevant notion of an "almost unitary" representation.

**Definition 2.7.**
1. *Let $\sigma$ be a continuous linear representation of $\Gamma$ on a Banach space $\mathcal{B}$. Given $\varepsilon > 0$, we say that $\sigma$ is $(K, \varepsilon)$-*almost unitary *if for any $k$ in $K$, the map $\sigma(k)$ is an $\varepsilon$-almost isometry.*
2. *If $\sigma$ is an $(\infty, s, \varepsilon, 0, 0)$-almost action of $\Gamma$ on a Banach space $\mathcal{B}$, we call $\sigma$ a $(K, \varepsilon, s)$-*almost unitary representation.

**Remark:** When a fixed choice of $K$ has been made, we frequently refer to a $(K, \varepsilon)$-almost unitary representation as an $\varepsilon$-almost unitary representation.

We begin by noting some consequences of Theorem 2.6 for a $(K, \varepsilon, s)$-almost unitary representation $\sigma$ of $\Gamma$ on a generalized $L^p$ space $\mathcal{B}$ where $1 < p < 2$ where $K, \varepsilon, s$ are chosen to satisfy the conclusions of Theorem 2.6 for some values of $M, C, h$. It is immediate that $\|\sigma(h)\| \leq (1+\varepsilon)^M$ and that $\|\sigma(h)^n\| \leq (1+\varepsilon)^{nM}$. We can define an operator $P$ by letting $Pv = \lim_{n\to\infty} \sigma(h)^n(v)$. It is easy to see that $\|\sigma(h)^{n+1} - \sigma(h)^n\| \leq C^n M$ and therefore that $\|\sigma(h)^n - P\| \leq C^{n-2} M$. One can then deduce that $\|P\| < 1 + \alpha$ where $\alpha$ depends only on $\varepsilon$ and $\alpha \to 0$ as $\varepsilon \to 0$.

If we have an $\varepsilon$-almost unitary representation $\sigma^*$ of $\Gamma$ on a generalized $L^q$ space $\mathcal{B}$ with $2 < q < \infty$, then the adjoint representation $\sigma$ of $\sigma^*$ on $\mathcal{B}^*$ is an $\varepsilon$-almost unitary representation of $\Gamma$ and $\mathcal{B}^*$ is a generalized

$L^p$ space for $1 < p < 2$. Assuming $f(\gamma) = f(\gamma^{-1})$ and therefore $h(\gamma) = h(\gamma^{-1})$, it follows that $\sigma(h)^* = \sigma^*(h)$. Since $\|A\| = \|A^*\|$ for any bounded operator $A$, so the estimates above carry over for $\sigma^*(h)$, and $\|\sigma^*(h)\|\leq(1+\varepsilon)^M$ and $\|\sigma^*(h)^n\|\leq(1+\varepsilon)^{nM}$. Furthermore, the operator $P^*$ defined by letting $P^*v = \lim_{n\to\infty}\sigma^*(h)^n(v)$ is the adjoint of $P$ and so bounded and a projection. Ideally, $P^*$ would project on $\Gamma$ invariant vectors. This is easy to verify if $\varepsilon = 0$, but unclear in general. It is also immediate that $\|\sigma^*(h)^{n+1} - \sigma^*(h)^n\|\leq C^nM$ and that $\|\sigma^*(h)^n - P^*\|\leq C^{n-2}M$. We summarize this discussion as follows:

**Corollary 2.8.** *If $\Gamma$ has property $(T)$ and $f\in\mathcal{U}_2(\Gamma)$ satisfies $f(\gamma) = f(\gamma^{-1})$ and $0 < C < 1$ and $1 < p_0 < \infty$, there exist positive integers $M = M(f,p_0), s = s(f,p_0)$ and $m = m(f,p_0)$ and $\varepsilon = \varepsilon(p_0) > 0$ such that, letting $h = f^{*m}$, for any $p < p_0$ and any $(K,\varepsilon,s)$-almost unitary representation $\sigma$ of $\Gamma$ on a generalized $L^p$ space $\mathcal{B}$ and any vector $v$, we have $d_{\mathcal{B}}(\sigma(h)^{n+1}v, \sigma(h)^n(v)) < MC^n\mathrm{disp}_K(v)$. Furthermore $Pv = \lim_{n\to\infty}\sigma(h)^n v$ is a bounded linear operator such that $d_{\mathcal{B}}(v,Pv)\leq\frac{MC}{1-C}\,\mathrm{disp}_K(v)$.*

**Remarks:**

(1) We emphasize again that we do not know if $Pv$ is necessarily $\Gamma$ invariant unless $\sigma$ is unitary. For applications, we will be dealing with Banach spaces $\mathcal{B}$ which are $L^p$ type function spaces and so subspaces of a Hilbert space $\mathcal{H}$ which is a function space of type $L^2$. The operator $\sigma(h)$ will be defined on $\mathcal{H}$ and we will know, by Theorem 2.1, that $\sigma(h)^n v$ converges to a $\Gamma$ invariant vector $v'$. Corollary 2.8 will be used in conjunction with the Sobolev embedding theorems to obtain stronger estimates on the regularity of $v'$. For this argument to work, it is important to know that we can choose $h$ satisfying both Corollary 2.8 and Theorem 2.1 at the same time. It is for this reason that we emphasize throughout that $h$ can be any large enough convolution power of any $f\in\mathcal{U}_2(\Gamma)$.

(2) We only explicitly use below the variant of this corollary for $(K,\varepsilon)$-almost unitary representations. As remarked above, the version for partially defined representations can be used in conjunction with Theorem 2.5 to give a proof of Theorem 2.11 that does not use Theorem 2.4.

2.3. **Foliating Theorem 1.1.** We now discuss the necessary notions to state our generalization of Theorem 1.1. Though our applications are to smooth foliations of smooth manifolds, here we work in a broader setting.

To motivate the results in this section, we state one corollary of the results of [FM2], for which Theorem 2.11 is a key ingredient in the proof. We call an action $\rho$ of a group $\Gamma$ on $\mathbb{T}^n$ *linear* if it is defined by a homomorphism from $\Gamma$ to $GL(n,\mathbb{Z})$, the full group of linear automorphisms of $\mathbb{T}^n$.

**Corollary 2.9** ([FM2]). *Let $G$ be a semisimple Lie group with all simple factors of real rank at least $2$ and let $\Gamma < G$ be a lattice. Then any linear action of $\Gamma$ on $\mathbb{T}^n$ is $C^{\infty,\infty}$ locally rigid and there exists a positive integer $k_0$ depending on the action, such that the action is $C^{k,k-\frac{n}{2}-2}$ locally rigid for all $k{\geq}k_0$.*

This result follows from a more general local rigidity theorem in [FM2] whose proof uses both Theorem 2.11 and our results from [FM1].

Throughout this section $X$ will be a locally compact, second countable metric space and $\mathfrak{F}$ will be a foliation of $X$ by $n$ dimensional manifolds. For background on foliated spaces, their tangent bundles, and transverse invariant measures, the reader is referred to [CC] or [MS]. Recall that $\mathfrak{F}$ is a partition of $X$, satisfying certain additional conditions, into smooth manifolds called *leaves* of the foliation. We will often refer to the leaf containing $x$ as $\mathfrak{L}_x$.

We let $\mathrm{Diff}^k(X,\mathfrak{F})$ be the group of homeomorphisms of $X$ which preserve $\mathfrak{F}$ and restrict to $C^k$ diffeomorphisms on each leaf with derivatives depending continuously on $x$ in $X$. For $1{\leq}k{\leq}\infty$, there is a natural $C^k$ topology on $\mathrm{Diff}^k(X,\mathfrak{F})$. The definition of this topology is straightforward and we sketch it briefly. As is usual, the topology on $\mathrm{Diff}^\infty(X,\mathfrak{F})$ is the inverse limit of the topologies on $\mathrm{Diff}^k(X,\mathfrak{F})$ so we now restrict to the case of $k$ finite. If $X$ is compact, we fix a finite cover of $X$ by charts $\tilde{U}_i$ which are products, such that there are proper subsets $U_i{\subset}\tilde{U}_i$ which are also products and which cover $X$. Without loss of generality, we can identify each $U_i$ as $B(0,r){\times}V_i$ where $B(0,r)$ is standard Euclidean ball and $V_i$ is an open set in the transversal and identify $\tilde{U}_i$ as $B(0,2r){\times}\tilde{V}_i$ where $\tilde{V}_i$ is an open set in the transversal such that $V_i{\subsetneq}\tilde{V}_i$. (See Proposition 7.2 below for a precise description of such charts.) A neighborhood of the identity in $\mathrm{Diff}^k(X,\mathfrak{F})$ will consist of homeomorphisms $\phi$ which map each $U_i$ inside $\tilde{U}_i$ and which are uniformly $C^k$ small as maps from each $B(0,r){\times}\{v\}$ to $B(0,2r){\times}\{v'\}$, where $v'$ is the point in $\tilde{V}_i$ such that $\phi(B(0,r){\times}\{v\}){\subset}B(0,2r){\times}\{v'\}$. When $X$ is noncompact, there are two possible topologies on $\mathrm{Diff}^k(X,\mathfrak{F})$. The *weak topology* is given by taking the inverse limit of the topologies described above for an increasing union of compact subsets of $X$. To define the *strong topology*, we cover $X$ by a countable collection of neighborhoods

$U_i \subset \tilde{U}_i$ as described above, and take the same topology. When $X$ is not compact, we will always consider the strong topology. Though non-compact foliated spaces arise in the proofs, for the remainder of this subsection, we consider only compact $X$.

We now define the type of perturbations of actions that we will consider.

**Definition 2.10.** *Let $\Gamma$ be a compactly generated topological group and $\rho$ an action of $\Gamma$ on $X$ defined by a homomorphism from $\Gamma$ to $\mathrm{Diff}^\infty(X, \mathfrak{F})$. Let $\rho'$ be another action of $\Gamma$ on $X$ defined by a homomorphism form $\Gamma$ to $\mathrm{Diff}^k(X, \mathfrak{F})$. Let $U$ be a (small) neighborhood of the identity in $\mathrm{Diff}^k(X, \mathfrak{F})$ and $K$ be a compact generating set for $\Gamma$. We call $\rho'$ a $(U, C^k)$-*foliated perturbation *of $\rho$ if:*

(1) *for every leaf $\mathfrak{L}$ of $\mathfrak{F}$ and every $\gamma \in \Gamma$, we have $\rho(\gamma)\mathfrak{L} = \rho'(\gamma)\mathfrak{L}$ and,*
(2) *$\rho'(\gamma)\rho(\gamma)^{-1}$ is in $U$ for every $\gamma$ in $K$.*

We fix a continuous, leafwise smooth Riemannian metric $g_{\mathfrak{F}}$ on $T\mathfrak{F}$, the tangent bundle to the foliation and note that $g_{\mathfrak{F}}$ defines a volume form and corresponding measure on each leaf $\mathfrak{L}$ of $\mathfrak{F}$, both of which we denote by $\nu_{\mathfrak{F}}$. (Metrics $g_{\mathfrak{F}}$ exist by a standard partition of unity argument.) Let $\Gamma$ be a group and $\rho$ an action of $\Gamma$ on $X$ defined by a homomorphism from $\Gamma$ to $\mathrm{Diff}^k(X, \mathfrak{F})$. We say the action is *leafwise isometric* if $g_{\mathfrak{F}}$ is invariant under the action. When $\Gamma = \mathbb{Z}$ and $\mathbb{Z} =< f >$, we will call $f$ a *leafwise isometry.*

For the remainder of the paper, we will assume that the foliation has a transverse invariant measure $\nu$. By integrating the transverse invariant measure $\nu$ against the Riemannian measure on the leaves of $\mathfrak{F}$, we obtain a measure $\mu$ on $X$ which is finite when $X$ is compact. In this case, we normalize $g_{\mathfrak{F}}$ so that $\mu(X) = 1$. We will write $(X, \mathfrak{F}, g_{\mathfrak{F}}, \mu)$ for our space equipped with the above data, sometime leaving one or more of $\mathfrak{F}, g_{\mathfrak{F}}$ and $\mu$ implicit. We will refer to the subgroup of $\mathrm{Diff}^k(X, \mathfrak{F})$ which preserves $\nu$ as $\mathrm{Diff}^k_\nu(X, \mathfrak{F})$. Note that if $\rho$ is an action of $\Gamma$ on $X$ defined by a homomorphism into $\mathrm{Diff}^k_\nu(X, \mathfrak{F})$ and $\rho$ is leafwise isometric, then $\rho$ preserves $\mu$. Furthermore if $\rho$ is an action of $\Gamma$ on $X$ defined by a homomorphism into $\mathrm{Diff}^k_\nu(X, \mathfrak{F})$ and $\rho'$ is a $(U, C^k)$-leafwise perturbation of $\rho$, then it follows easily from the definition that $\rho'$ is defined by a homomorphism into $\mathrm{Diff}^k_\nu(X, \mathfrak{F})$ since the induced map on transversals is the same.

The following foliated version of Theorem 1.1 is one of the key steps in the proof of the main results in [FM2]. We denote by $B_{\mathfrak{F}}(x, r)$ the ball in $\mathfrak{L}_x$ about $x$ of radius $r$. For a sufficiently small value of $r > 0$,

we can canonically identify each $B_{\mathfrak{F}}(x,2r)$ with the ball of radius $2r$ in Euclidean space via the exponential map from $T\mathfrak{F}_x$ to $\mathfrak{L}_x$. To state our results, we will need a quantitative measure of the size of the $k$-jet of $C^k$ maps. We first consider the case when $k$ is an integer, where we can give a pointwise measure of size. Recall that a $C^k$ self map of a manifold $Z$ acts on $k$-jets of $C^k$ functions on $Z$. Any metric on $TZ$ defines a pointwise norm on each fiber of the bundle of $J^k(Z)$ of $k$-jets of functions on $Z$. For any $C^k$ diffeomorphism $f$ we can define $\|j^k(f)(z)\|$ as the operator norm of the map induced by $f$ from $J^k(Z)_z$ to $J^k(Z)_{f(z)}$. For a more detailed discussion on jets and an explicit construction of the norm on $J^k(Z)_z$, see section 4. We say that a map $f$ has $C^k$ size less than $\delta$ on a set $U$ if $\|j^k(f)(z)\| < \delta$ for all $z$ in $U$. If $k$ is not an integer, we say that $f$ has $C^k$ size less than $\delta$ on $U$ if $f$ has $C^{k'}$ size less than $\delta$ on $U$ where $k'$ is the greatest integer less than $k$ and $j^{k'}(f)$ satisfies a (local) Hölder estimate on $U$. See section 4 for a more detailed discussion of Hölder estimates.

**Remark:** This notion of $C^k$ size is not very sharp. The size of the identity map will be 1, as will be the size of any leafwise isometry. We only use this notion of size to control estimates on a map at points where the map is known to be "fairly large" and where we only want bounds to show it is "not too large".

For the following theorem, we assume that the holonomy groupoid of $(X,\mathfrak{F})$ is Hausdorff. This is a standard technical assumption that allows us to define certain function spaces on "pairs of points on the same leaf of $(X,\mathfrak{F})$". See subsection 7.1, [CC] and [MS] for further discussion. All the foliations considered in [FM2] are covered by fiber bundles, in which case it is easy to show that the holonomy groupoid is Hausdorff.

**Theorem 2.11.** *Let $\Gamma$ be a locally compact, $\sigma$-compact group with property $(T)$. Let $\rho$ be a continuous leafwise isometric action of $\Gamma$ on $X$ defined by a homomorphism from $\Gamma$ to $\mathrm{Diff}^{\infty}_{\nu}(X,\mathfrak{F})$. Then for any $k{\geq}3, \kappa>0$ and any $\varsigma>0$ there exists a neighborhood $U$ of the identity in $\mathrm{Diff}^k(X,\mathfrak{F})$ such that for any continuous $(U,C^k)$-foliated perturbation $\rho'$ of $\rho$ there exists a measurable $\Gamma$-equivariant map $\phi: X{\to}X$ such that:*

1. *$\phi{\circ}\rho(\gamma)=\rho'(\gamma){\circ}\phi$ for all $\gamma{\in}\Gamma$,*
2. *$\phi$ maps each leaf of $\mathfrak{F}$ into itself,*
3. *there is a subset $S{\subset}X$ with $\mu(S)=1-\varsigma$ and $\Gamma{\cdot}S$ has full measure in $X$, and a constant $r{\in}\mathbb{R}^+$, depending only on $X,\mathfrak{F}$ and $g_{\mathfrak{F}}$, such that, for every $x{\in}S$, the map $\phi: B_{\mathfrak{F}}(x,r){\to}\mathfrak{L}_x$ is $C^{k-1-\kappa}$-close to the identity; more precisely, with our chosen*

*identification of $B_{\mathfrak{F}}(x,2r)$ with the ball or radius $2r$ in Euclidean space, $\phi - \mathrm{Id} : B_{\mathfrak{F}}(x,r) \rightarrow B_{\mathfrak{F}}(x,2r)$ has $C^{k-1-\kappa}$ norm less than $\varsigma$ for every $x \in S$, and*

(4) *there exists $0 < t < 1$ depending only on $\Gamma$ and $K$ such that the set of $x \in X$ where the $C^{k-1-\kappa}$ size of $\phi$ on $B_{\mathfrak{F}}(x,r)$ is not less than $(1+\varsigma)^{l+1}$ has measure less than $t^l \varsigma$ for any positive integer $l$.*

*Furthermore, for any $l \geq k$, if $\rho'$ is a $C^{2l-k+1}$ action, then by choosing $U$ small enough, we can choose $\phi$ to be $C^l$ on $B_{\mathfrak{F}}(x,r)$ for almost every $x$ in $X$. In particular, if $\rho'$ is $C^\infty$ then for any $l \geq k$, by choosing $U$ small enough, we can choose $\phi$ to be $C^l$ on $B_{\mathfrak{F}}(x,r)$ for almost every $x$ in $X$.*

**Remarks:**

(1) The map $\phi$ constructed in the theorem is not even $C^0$ close to the identity on $X$. However, the proof of the theorem shows that for every $1 \leq q < \infty$, possibly after changing $U$ depending on $q$, we have $\int_X (d(x,\phi(x))^q d\mu \leq \varsigma$.

(2) This theorem implies a version of Theorem 1.1, but with lower regularity.

(3) In some special cases it is possible to slightly improve the regularity of $\phi$. It is possible to show that $\phi$ is $C^l$ for some given choice of $l$ even if $\rho'$ is only $C^{l+1}$ provided $U$ is small enough, see section 6 for more discussion. Unlike in Theorem 1.1, it does not seem possible to show that $\phi$ is $C^\infty$ without some assumption on the action transverse to $\mathfrak{F}$. Again see section 6 for more details.

In the case when $X$ is a direct product, we can prove slightly greater regularity.

**Theorem 2.12.** *If $X = Y \times Z$ and the foliation $\mathfrak{F}$ has leaves of the form $\{\{y\} \times Z | y \in Y\}$, then $\phi$ in Theorem 2.11 is $C^{k-\kappa}$ and all estimates in that theorem for the $C^{k-1-\kappa}$ topology can be replaced by analogous estimates in the $C^{k-\kappa}$ topology.*

**Remark:** We do not give a proof of Theorem 2.12 here. The proof of Theorem 1.1 given in subsection 5.1 can be combined with the techniques of section 7 to give such a proof, which we leave as an exercise for the interested reader.

## 3. **Proof of Theorem 2.5 and variants**

In this section we prove Theorem 2.5. In the first subsection, we give a proof of the analogue of Theorem 2.5 for isometric actions of groups

with property $(T)$ on Hilbert spaces. In the second subsection we develop a general method of constructing limit actions from sequences of actions. In the third subsection, we prove Theorem 2.5 modulo some observations contained in the appendix to this paper, which are required only when the action is not affine and $\Gamma$ is not discrete. In the final subsection we recall some results from [BFGM] and some facts about Banach spaces of type $L^p_n$ and indicate the modifications to prior arguments needed to prove the results in subsection 2.2.

3.1. **Finding fixed points for isometric actions of groups with Property** $(T)$**.** Theorem 2.5 is a generalization of the following consequence of property $(T)$. Though this fact is a variant of well-known consequences of property $(T)$, we did not find a prior reference for this precise statement and so give a detailed proof.

We first fix some notation. As in subsection 2.1 we fix a locally compact, $\sigma$-compact group $\Gamma$ with a compact generating set $K{\subset}\Gamma$ containing a neighborhood of the identity, and a (left) Haar measure $\mu$ on $\Gamma$. Given a function $h{\in}C_c(\Gamma)$ and $\gamma_0{\in}\Gamma$ we write $\gamma_0{\cdot}h$ for the function $\gamma{\to}h({\gamma_0}^{-1}\gamma)$. The subsets $\mathcal{U}(\Gamma)$ and $\mathcal{U}_2(\Gamma)$ of $C_c(\Gamma)$ are as in subsection 2.1

**Proposition 3.1.** *If $\Gamma$ has property $(T)$ and $f{\in}\mathcal{U}_2(\Gamma)$ and $0 < C_0 < 1$ there exist positive integers $M = M(f, C_0)$ and $m = m(f, C_0)$, such that, letting $h = f^{*m}$, for any Hilbert space $\mathcal{H}$, any continuous isometric action $\rho$ of $\Gamma$ on $\mathcal{H}$, and any $x{\in}\mathcal{H}$ we have*

(1) $d_{\mathcal{H}}(x, \rho(h)(x)){\leq}M\,\mathrm{disp}_K(x)$
(2) $\mathrm{disp}_K(\rho(h)(x)){\leq}C_0\,\mathrm{disp}_K(x)$.

Given a unitary representation $\sigma$ of $\Gamma$ on $\mathcal{H}$, we let $\mathcal{H}_\sigma$ be the $\sigma$ invariant vectors and $\mathcal{H}_\sigma^\perp$ it's orthogonal complement. We recall a fact about groups with property $(T)$.

**Lemma 3.2.** *Let $\mathcal{H}$ be a Hilbert space and $\sigma$ a continuous unitary representation of $\Gamma$ on $\mathcal{H}$. Then for any $f{\in}\mathcal{U}_2(\Gamma)$, we have $\sigma(f)|_{\mathcal{H}_\sigma^\perp}$ is a contraction. More precisely, there exists a constant $0 < D < 1$ such that $\|\sigma(f)(x)\| < D\|x\|$ for any $x{\in}\mathcal{H}_\sigma^\perp$.*

This lemma is an immediate consequence of Kazhdan's definition of property $(T)$ and the characterization of the Fell topology in Lemma III.1.1 of [M]. Though explicitly stated there only for some $f$, the proof is valid for any $f{\in}\mathcal{U}_2(\Gamma)$. For a proof of a more general fact see Lemma 3.16 below. The following lemma is elementary from the fact that isometries of Hilbert spaces are affine [MU].

**Lemma 3.3.** *Let $\Gamma$ be a group, $\mathcal{H}$ a Hilbert space and $\rho$ an isometric $\Gamma$ action on $\mathcal{H}$. Then for any measures $\mu, \lambda \in \mathcal{P}(\Gamma)$, we have $\rho(\mu)\rho(\lambda) = \rho(\mu * \lambda)$.*

**Lemma 3.4.** *If $\Gamma$ has property $(T)$ and $f \in \mathcal{U}(\Gamma)$ and $0 < C_0 < 1$, there exist positive integers $M = M(f, C_0)$ and $m = m(f, C_0)$ such that, letting $h = f^{*m}$, for any Hilbert space $\mathcal{H}$, any continuous unitary representation $\sigma$ of $\Gamma$ on $\mathcal{H}$, and any $v \in \mathcal{H}$ we have*

(1) $d_{\mathcal{H}}(v, \sigma(h)(v)) \leq M \operatorname{disp}_K(v)$
(2) $\operatorname{disp}_K(\sigma(h)(v)) \leq C_0 \operatorname{disp}_K(v)$.

*Proof.* Since if $v = (v_1, v_2)$, where $v_1 \in \mathcal{H}_\sigma$ and $v_2 \in \mathcal{H}_\sigma^\perp$, $\operatorname{disp}_K(v) = \operatorname{disp}_K(v_2)$, it suffices to assume $\mathcal{H} = \mathcal{H}_\sigma^\perp$. Since $\Gamma$ has property $(T)$ this implies that there exists $\varepsilon$ such that there are no $(K, \varepsilon)$-invariant vectors in $\mathcal{H}$, i.e.

$$\varepsilon\|v\| < \|\sigma(k)v - v\| \leq 2\|v\|$$

for any $k \in K$ and any $v \in \mathcal{H}$. Let $D$ be the contraction factor from 3.2 and choose $m$ such that $2D^m \leq C_0\varepsilon$ and let $h = f^{*m}$. Note that Lemma 3.3 implies that $\sigma(h) = \sigma(f^{*m}) = \sigma(f)^m$. Let $v$ be a vector with $K$-displacement $\delta$. It follows from the equation above that $\delta > \varepsilon\|v\|$. Direct computation shows that $\|\sigma(k)\sigma(h)(v) - \sigma(h)(v)\| < 2\|\sigma(h)v\| < 2D^m\|v\| \leq C_0\varepsilon\|v\| < C_0\delta$ which is the second conclusion of the lemma. Letting $M$ be the smallest value such that $\operatorname{supp}(h) \subset K^M$, the first conclusion follows as well. □

*Proof of Proposition 3.1.* Fix the function $h \in \mathcal{U}(\Gamma)$ and the constant $0 < C_0 < 1$ from the conclusion of Lemma 3.4. As any continuous affine, isometric action of a group with property $(T)$ on a Hilbert space has a fixed point, $\Gamma$ fixes some point $x$ in $\mathcal{H}$ [De]. Viewing $\mathcal{H}$ as a vector space with $x$ as origin allows us to view our action $\rho$ as a unitary representation, and the proposition is now an immediate consequence of Proposition 3.4 with the same $h, M$ and $C_0$.

□

3.2. **Limits of sequences of actions.** In this subsection we give a very general process for constructing a limit action from a sequence of actions, or partially defined actions. The reader primarily interested in actions of discrete groups may compare this with the discussion of scaling limits in [Gr2] and the references cited there. Throughout this subsection $\Gamma$ is a group and $K$ is a generating set for $\Gamma$.

Let $X_n$ be a sequence of complete metric spaces, with distinguished points $x_n \in X_n$, and let $\rho_n$ be $(r_n, s_n, \varepsilon_n, \delta_n, K)$-almost actions of a group $\Gamma$ on $X_n$ at $x_n$. We construct our limit space $X$ as a quotient of a certain

subspace $\tilde{X}$ in $\prod X_n$. We will use ultrafilters and ultralimits to define $\tilde{X}$, and a pseudo-metric on $\tilde{X}$, and then let $X$ be $\tilde{X}$ modulo relation of being at distance zero in the pseudo-metric.

**Definition 3.5.** *A* non-principal ultrafilter *is a finitely additive probability measure $\omega$ defined on all subsets of $\mathbb{N}$ such that*

(1) $\omega(S) = 0$ *or* $1$,
(2) $\omega(S) = 0$ *if $S$ is finite.*

This definition is the one given by Gromov in [Gr1], at the beginning of section 2.$A$. on page 36. It is not clear with this definition that non-principal ultrafilters exist. To show existence, one defines an *ultrafilter* as a maximal *filter*, and shows that maximal objects exist using Zorn's lemma. For a more traditional definition of ultrafilters, see [BTG, I.6.4]. In the context of group theory ultrafilters were first used to construct limits of sequences of metric spaces in [vDW], though their use in the study of Banach space theory is much older than that, see [BL] and [He] for more history. In what follows, we fix a non-principal ultrafilter $\omega$.

Let $\{y_n\}$ be sequence in $\mathbb{R}$, the *$\omega$-limit* of $\{y_n\}$ is $\omega\text{-}\lim y_n = y$ if for every $\epsilon > 0$ it follows that $\omega\{n|d(y_n, y) < \epsilon\} = 1$. The following well-known proposition can be proven easily by mimicking the proof that bounded sequences have limit points.

**Proposition 3.6.** *Any bounded sequence of real numbers has a unique $\omega$-limit.*

More generally, if $X$ is a Hausdorff topological space, and $\{y_n\}$ is a sequence of points in $X$, the *$\omega$-limit* of $\{y_n\}$ is $\omega\text{-}\lim y_n = y$ if every neighborhood of $y$ has measure 1 with respect to the pushforward of $\omega$ under the map $n \rightarrow y_n$. The following, almost tautological proposition, is from [BTG, I.10.1]:

**Proposition 3.7.** *The space $X$ is compact if and only if, for every ultrafilter $\omega$, every sequence $\{y_n\}$ has a unique $\omega$-limit.*

We let $\tilde{X} = \{y \in \prod X_n | \omega\text{-}\lim d_n(y_n, x_n) < \infty\}$. We put a metric on $\tilde{X}$ by letting $\tilde{d}(\{v_n\}, \{w_n\}) = \omega\text{-}\lim d_n(v_n, w_n)$. It is easy to check that $\tilde{d}$ is a pseudo-metric on $\tilde{X}$. We can define an equivalence relation on $\tilde{X}$ by letting $v \sim w$ if $\tilde{d}(v, w) = 0$. We let $X = \tilde{X}/\sim$ with the metric $d$ defined by $\tilde{d}$. For an arbitrary sequence $y \in \tilde{X}$, we refer to the image of $y = \{y_n\}$ in $X$ as $y_\omega$ and write $y_\omega = \omega\text{-}\lim y_n$. The space $X$ has a natural basepoint given by $x_\omega = \omega\text{-}\lim x_n$. The space $(X, d)$ is often called *the $\omega$-ultraproduct*, or simply *the ultraproduct*, with $\omega$ implicit, of $(X_i, d_i, x_i)$. The following straightforward proposition is standard.

**Proposition 3.8.** *The space $(X,d)$ is complete.*

*Proof.* Let $x^j_\omega$ be a Cauchy sequence in $X$, where $x^j_\omega = \omega\text{-}\lim x^j_n$. Let $N_1 = \mathbb{N}$. Inductively, there is an $\omega$-full measure subset $N_j \subseteq N_{j-1}$ such that $n \in N_j$ implies that $|d_n(x^k_n, x^l_n) - d(x^k_\omega, x^l_\omega)| \leq \frac{1}{2^n}$ for $1 \leq k, l \leq j$. For $n \in N_j \backslash N_{j-1}$, define $y_n = x^j_n$. Then $x^j_\omega$ converge to $y_\omega$. $\square$

We record here one additional fact about limits of sequences of Hilbert spaces that we will use in the proof of Theorem 2.5, compare [He].

**Proposition 3.9.** *If the spaces $X_n$ are Hilbert spaces with $x_n = 0$ and inner product $<,>_n$, then the space $(X,d)$ is Hilbert space with $x_\omega = 0$ and inner product defined by $< v_\omega, w_\omega >= \omega\text{-}\lim < v_n, w_n >$.*

*Proof.* Since we already know $X$ is complete, we need only check that $X$ is a vector space and that $< \cdot, \cdot >$ is a positive definite symmetric bilinear form. Letting $V = \{v \in \tilde{X} | d(v, \{x_n\}) = 0\}$, it is immediate that $V$ is a sub-vector space of the vector space $\tilde{X}$ and that $\tilde{X}/V = \tilde{X}/\sim$.

That $<,>$ is symmetric and bilinear is immediate. The definition implies that $< v_\omega, v_\omega >= d(v_\omega, x_\omega)^2 > 0$ if $v_\omega \neq x_\omega$, so the form is positive definite. $\square$

We now proceed to define a $\Gamma$ action $\rho$ on $X$. If $\delta_n$ and $\varepsilon_n$ were bounded sequences and each $\rho_n$ were globally defined, we could define a $\Gamma$ action $\tilde{\rho}$ on $\tilde{X}$ simply by acting on each coordinate. Instead we define $\tilde{\rho}(\gamma)(y)$ to be the sequence whose $n$th coordinate is $\rho_n(\gamma)(y_n)$ when $\rho_n(\gamma)(y_n)$ is defined and whose $n$th coordinate is $x_n$ otherwise. Though this is not an action, we have:

**Proposition 3.10.** *If $\omega\text{-}\lim \varepsilon_n = \varepsilon < \infty, \omega\text{-}\lim \delta_n = \delta < \infty$ and $\omega\text{-}\lim_{n\to\infty} r_n = \omega\text{-}\lim_{n\to\infty} s_n = \infty$, then for every $\gamma$ in $\Gamma$, the map $\tilde{\rho}(\gamma)$ descends to a well-defined bilipschitz map $\rho(\gamma)$ of $X$ and the map $\rho : \Gamma \times X \to X$ is an action of $\Gamma$ on $X$. Furthermore, $\rho(k)$ is an $\varepsilon$-almost isometry of $X$ for every $k$ in $K$ and the $K$-displacement of $x_\omega$ is at most $\delta$.*

**Remark:** For our applications, we will have $\delta_n$ uniformly bounded, $\lim_{n\to\infty} \varepsilon_n = 0$, $\lim_{n\to\infty} r_n = \infty$ and $\lim_{n\to\infty} s_n = \infty$.

*Proof.* To verify that $\tilde{\rho}(\gamma)$ descends to $X$ and that it is bilipschitz, it suffices to verify that $\tilde{\rho}(\gamma)$ is an almost isometry of the pseudo-metric $\tilde{d}$. Let $u, v \in \tilde{X}$ and $\gamma \in \Gamma$. We fix the minimal $s$ such that $\gamma \in K^s$. By ignoring an $\omega$-measure zero, finite set of indices, we may assume $\rho_n(\gamma)u_n$ and $\rho_n(\gamma)v_n$ are defined. By definition

$$\tilde{d}(\tilde{\rho}(\gamma)u, \tilde{\rho}(\gamma)v) = \omega\text{-}\lim d_n(\rho_n(\gamma)u_n, \rho_n(\gamma)v_n).$$

Since $\rho_n(k)$ acts by $\varepsilon_n$-almost isometries for all $k\in K$, we have

$$(1-\varepsilon_n)^s d_n(u_n, v_n) \leq d_n(\rho_n(\gamma)u_n, \rho_n(\gamma)v_n) \leq (1+\varepsilon_n)^s d_n(u_n, v_n).$$

By taking the $\omega$-limit of the above equation, we have

$$(1-\varepsilon)^s \tilde{d}(u, v) \leq \tilde{d}(\tilde{\rho}(\gamma)u, \tilde{\rho}(\gamma)v) \leq (1+\varepsilon)^s \tilde{d}(u_n, v_n).$$

Which shows that $\tilde{\rho}(\gamma)$ preserve the equivalence relation of being at $\tilde{d}$ distance zero as well as showing that the map $\rho(\gamma)$ on the quotient $X$ is bilipschitz and in fact an $\varepsilon$-almost isometry when $\gamma\in K$.

That these maps form a $\Gamma$ action is almost obvious. Fix $\gamma_1, \gamma_2 \in \Gamma$ and $v\in\tilde{X}$. By ignoring a finite set $S_{\gamma_1,\gamma_2,v}$ of $\omega$-measure zero, we can insure that $\rho_n(\gamma_1\gamma_2)v_n, \rho_n(\gamma_2)v_n$ and $\rho_n(\gamma_1)(\rho_n(\gamma_2)v_n)$ are well-defined and that $\rho_n(\gamma_1\gamma_2)v_n = \rho_n(\gamma_1)(\rho_n(\gamma_2)v_n)$ for $n\notin S_{\gamma_1,\gamma_2,v}$. This implies $\rho(\gamma_1\gamma_2)v_\omega = \rho(\gamma_1)(\rho(\gamma_2)v_\omega)$. Since this verification (though not the set $S_{\gamma_1,\gamma_2,v}$) is independent of $\gamma_1, \gamma_2$ and $v$, it follows that $\rho$ is an action. That the $K$-displacement of $x_\omega$ is less than $\delta$ follows since $d(\rho(k)x_\omega, x_\omega) = \omega\text{-}\lim d_n(\rho_n(k)x_n, x_n) \leq \delta$ for all $k\in K$. $\square$

**Remark:** As shorthand for the construction above, we will write $\rho = \omega\text{-}\lim \rho_n$.

3.3. **Proof of Theorem 2.5.** In this subsection $\Gamma$ and $K$ are as in subsection 3.1. We fix the function $h\in\mathcal{U}(\Gamma)$ and the constant $C_0$ given by Proposition 3.1. We also fix an arbitrary $C$ with $C_0 < C < 1$.

*Proof of Theorem 2.5 for* $\Gamma$ *discrete.* Fix $\eta_0 = C - C_0$. The proof proceeds by contradiction, so we assume the theorem is false. Let $r_n = 2^n, s_n = n$ and $\varepsilon_n = \frac{1}{n}$ and $0 < \delta_n < \delta_0$. By the assumption that Theorem 2.5 is false there exists a sequence of Hilbert spaces $\mathcal{H}_n$, points $x_n\in\mathcal{H}_n$ and $(r_n, s_n, \varepsilon_n, \delta_n, K)$-almost actions $\rho_n$ at $x_n$ such that $\text{disp}_K(\rho_n(h)(x)) > C\delta_n$. By conjugation by a homothety at $x_n$, it suffices to consider the case where $\delta_n = 1$ for all $n$. Conjugating by this homothety makes $\rho_n$ a $(r_n\frac{1}{\delta_n}, s_n, \varepsilon_n, 1, K)$-almost action at $x_n$ and it remains true that $r_n\to\infty$ as $n\to\infty$. We will denote the distance on $\mathcal{H}_n$ as $d_n$ and the inner product as $< \cdot, \cdot >_n$. Letting $\tilde{\mathcal{H}}\subset\prod\mathcal{H}_n$ be as in the paragraph following Proposition 3.7 and $\tilde{d} = \omega\text{-}\lim d_n$ and $V$ the set of points in $\tilde{\mathcal{H}}$ with $\tilde{d}(v, \{x_n\}) = 0$ as above, it follows from Proposition 3.10, the fact that $\lim_{n\to\infty}\varepsilon_n = 0$, and Proposition 3.9 that the action $\rho = \omega\text{-}\lim\rho_n$ of $\Gamma$ on $\mathcal{H} = \tilde{\mathcal{H}}/V$ is an isometric action on a Hilbert space. It also follows from Proposition 3.10 that $\text{disp}_K(x_\omega)\leq 1$. By Proposition 3.1, this implies that $\text{disp}_K(\rho(h)(x_\omega))\leq C_0$. It is immediate from

the definitions that $\rho(h)(x_\omega) = \{\rho_n(h)x_n\}_\omega$ and that $d(\rho(k)y_\omega, y_\omega) = \omega$-$\lim d_n(\rho_n(k)y, y)$ for any $k \in K$ and $y \in \tilde{\mathcal{H}}$. Letting $y = \rho(h)(x_\omega)$, we have a set $S_k$ of full $\omega$-measure such that $d_n(\rho_n(k)\rho_n(h)(x_\omega), \rho_n(h)(x_\omega)) \leq C_0 + \eta_0$ for all $n$ in $S_k$. Letting $S = \cap_{k\in K} S_k$ we see that $\mathrm{disp}_K(\rho_n(h)(x_n)) \leq C = C_0 + \eta_0$ for any $n \in S$. Since $K$ is finite, $\omega(S) = 1$, and we have a contradiction. $\square$

Before proving the theorem for more general groups, we state some additional results needed because the limit action we construct is not necessarily continuous.

For the remainder of this subsection, we assume that each $\rho_n$ is continuous and that $\rho_n$ satisfy the conditions of Proposition 3.10. We can then define a limit action $\rho = \omega$-$\lim \rho_n$ as in Proposition 3.10. In general it is not true that $\rho$ is continuous, but we now describe a (possibly trivial) continuous subaction of $\rho$.

Given $y_n \in X_n$, we have an orbit map $\rho_n^{y_n} : K^{s^n} \to X_n$ defined by $\rho_n^{y_n}(\gamma) = \rho_n(\gamma)(y_n)$. We call a sequence $\{y_n\}$ *$\omega$-equicontinuous on compact sets* if for any compact subset $D$ of $\Gamma$, there exists a subset $S \subset \mathbb{N}$ with $\omega(S) = 1$ such that the orbit maps $\rho_n^{y_n}$ are equicontinuous on $D$ for $n$ in $S$. Since the collection of actions $\rho_n$ are uniformly bilipschitz, to prove a sequence is $\omega$-equicontinuous on compact sets, it suffices to prove that it is $\omega$-equicontinuous at the identity in $\Gamma$, i.e. given $\varepsilon > 0$, there is a neighborhood $U$ of the identity in $\Gamma$ such that $\rho_n^{y_n}(U) \subset B(y_n, \varepsilon)$ for $n$ in a set $S$ with $\omega(S) = 1$. We denote by $\Omega$ the set of $\omega$-equicontinuous sequences in $\tilde{X}$ and by $\bar{\Omega}$ the image of $\Omega$ in $X$. Keeping in mind that $\rho$ is an action by bilipschitz maps, it is straightforward to verify the following.

**Proposition 3.11.** *The set $\bar{\Omega}$ is closed and $\Gamma$ invariant. The restriction of $\rho$ to $\bar{\Omega}$ is continuous.*

We state a result giving sufficient conditions for $\bar{\Omega}$ to be an affine Hilbert subspace when the $X_n$ are all Hilbert spaces.

**Proposition 3.12.** *Let $\mathcal{H}_n$ be a sequence of Hilbert spaces with basepoints $x_n$. Let $\rho_n$ be a sequence of continuous $(r_n, s_n, \varepsilon_n, \delta_n, K)$-almost actions of $\Gamma$ on $X_n$ at $x_n$ with $\omega$-$\lim \varepsilon_n = 0, \omega$-$\lim \delta_n = \delta < \infty$ and $\omega$-$\lim_{n\to\infty} r_n = \omega$-$\lim_{n\to\infty} s_n = \infty$. Then the set $\bar{\Omega} \subset \mathcal{H}$ is an affine Hilbert subspace of $\mathcal{H}$. Furthermore if $f \in \mathcal{U}(\Gamma)$, then for any sequence $\{y_n\} \in \tilde{\mathcal{H}}$, the point $\omega$-$\lim \rho_n(f)y_n$ is in $\bar{\Omega}$.*

We will also need the following generalization of Lemma 3.3 for almost isometric actions on Hilbert spaces.

**Proposition 3.13.** *Let $\mathcal{H}$ be a Hilbert space. Then for every $r, \eta > 0$ there is a $\varepsilon > 0$ such that for any continuous $(r, s, \varepsilon, \delta, K)$-almost action of $\Gamma$ on $\mathcal{H}$ at $x$ and any measures $\mu, \lambda \in \mathcal{P}(\Gamma)$ with $\operatorname{supp}(\mu), \operatorname{supp}(\lambda)$ and $\operatorname{supp}(\mu * \lambda)$ contained in $K^s$ and $s\delta < \frac{r}{2}$, we have*

$$d(\rho(\mu)\rho(\lambda)x, \rho(\mu * \lambda)x) \leq \eta.$$

**Remark:** For our applications to local rigidity, it suffices to prove Theorem 2.5 for affine $\varepsilon$-almost isometric actions. In this case, the proof of Theorem 2.5 is almost the same, but we can assume $\rho_n$ is affine for all $n$. We will therefore only prove those cases of Propositions 3.12 and 3.13 here. Only readers interested in Theorem 2.5 for the case of $\rho$ not affine and $\Gamma$ not discrete, need refer to Appendix A of this paper for proofs of the general cases of Propositions 3.12 and 3.13.

*Proof of Propositions 3.12 and 3.13 for affine actions.* It is immediate that $\bar{\Omega}$ is an affine Hilbert subspace and that $\rho_n(\mu)\rho_n(\lambda) = \rho_n(\mu * \lambda)$. We now prove that $\rho_n(f)y_n$ is $\omega$ equicontinuous. To do so we use the following estimate:

$$\begin{gathered} d(\rho_n(\gamma_0)\rho_n(f)y_n, \rho_n(f)y_n) \leq \\ \|\rho(\gamma_0 \cdot f - f)y_n\| \leq \\ \|\gamma_0 \cdot f - f\|_{L^1} D_{\gamma_0, f} \end{gathered}$$

where $D_{\gamma_0,f} = \sup_{\operatorname{supp}(\gamma_0 \cdot f - f)} d(\rho_n(\gamma)x, x)$. This estimate, our assumptions on $\rho_n$, the fact that $K$ contains a neighborhood of the identity in $\Gamma$, and continuity of the $\Gamma$ action on $L^1(\Gamma)$ imply that for any $\eta > 0$ there is a neighborhood $U$ of the identity in $\Gamma$ such that whenever $\gamma_0 \in U$, we have $d(\rho_n(\gamma_0)\rho_n(f)y_n, \rho_n(f)y_n) \leq \eta$. □

**Remark:** The proof of the first assertion of Proposition 3.12 for non-affine actions occurs in subsection A.1 and the proof of the second assertion is found following the proof of Lemma A.3 in subsection A.2. The proof of Proposition 3.13 for non-affine actions is found at the end of subsection A.3.

In the discrete case, we implicitly used finiteness of $K$ to show that the $K$-displacement of the $\omega$-limit of a sequence is equal to the $\omega$-limit of the $K$-displacements. This is true more generally for sequences which are $\omega$-equicontinuous on compact sets.

**Proposition 3.14.** *Let $\{y_n\} \in \Omega$. Then $\operatorname{disp}_K(y_\omega) = \omega\text{-}\lim \operatorname{disp}_K(y_n)$.*

*Proof.* We let $k_n$ be the sequence of elements in $K$ such that the $K$-displacement of $y_n$ is $d(\rho_n(k_n)y_n, y_n)$. By Proposition 3.7, there is a unique $k_\omega = \omega\text{-}\lim k_n$. Since $y_n \in \Omega$, we know that $d(\rho_n(k)y_n, y_n)$ are

equicontinuous functions of $k \in K$. This implies $\omega\text{-}\lim d(\rho_n(k_\omega)y_n, y_n) = \omega\text{-}\lim d(\rho_n(k_n)y_n, y_n)$ which suffices to prove the proposition. $\square$

Lastly, we need the following trivial lemma.

**Lemma 3.15.** *Let $X$ be a metric space and $\rho$ an $(r, s, \varepsilon, \delta, K)$-almost action of $\Gamma$ on $X$ at $x$. Then if $d(x, y) = \eta$, then $\mathrm{disp}_K(y) \leq \delta + 2\eta + \varepsilon\eta$.*

*Proof of Theorem 2.5 for non-discrete $\Gamma$.* The proof begins as in the discrete case, we assume the theorem is false and let $\mathcal{H}_n, x_n$ be a sequence of Hilbert spaces and $\rho_n$ be as in the proof for $\Gamma$ discrete, assuming we have already renormalized so $\delta_n = 1$. We note that to prove the theorem it suffices to show the existence of some $h \in \mathcal{U}(\Gamma)$, so we may assume that $\mathrm{disp}_K(\rho(h)x_n) > C\,\mathrm{disp}_K(x_n)$ for every $h \in \mathcal{U}(\Gamma)$. Arguing as in the discrete case, we can produce a isometric limit action $\rho$ on Hilbert space $\mathcal{H}$ where the $K$-displacement of $x_\omega = \omega\text{-}\lim x_n$ is 1. By Proposition 3.12, $\rho$ is continuous on a closed affine subspace $\mathcal{H}' \subset \mathcal{H}$ and for any sequence $\{y_n\} \in \tilde{\mathcal{H}}$ and any $g \in \mathcal{U}(\Gamma)$, the point $\omega\text{-}\lim \rho_n(g)y_n$ is in $\mathcal{H}'$. Together with Proposition 3.14, the arguments for the discrete case imply that for any $\{y_n\}$ with $y_\omega \in \mathcal{H}'$, we have $\mathrm{disp}_K(\rho(h)y_n) \leq C\,\mathrm{disp}_K(y_n)$ for $\omega$-almost every $n$. If $x_\omega \in \mathcal{H}'$ this completes the proof. Otherwise let $x'_n = \rho(f)x_n$ where $f \in \mathcal{U}(\Gamma)$ with $\mathrm{supp}(f) \subset K$ and $\mathrm{supp}(f)$ containing a neighborhood of the identity in $\Gamma$.

In this case, we will prove the theorem with $h$ replaced by $h^{*d} * f$ for a positive integer $d$ such that $4C^d \leq C$. We know from Proposition 3.12 that the sequence $\{\rho_n(f)x_n\}$ is equicontinuous, as is $\{\rho_n(h)^i\rho_n(f)x_n\}$ for every positive integer $i$. Since $\rho_n(f)x_n$ is in the ball of radius 1 about $x$, Lemma 3.15 implies that $\mathrm{disp}_K(\rho_n(f)x_n) \leq 3 + \varepsilon_n$. Therefore, we know that $\mathrm{disp}_K(\rho_n(h)^d\rho_n(f)x_n) \leq C^d(3 + \varepsilon_n)$. Choosing $\eta < \frac{C^d}{10d}$ by $d$ applications of Proposition 3.13, we have that $d_n(\rho_n(h)^d\rho_n(f)x_n, \rho_n(h^{*d} * f)x_n) < \frac{C^d}{10}$ for $\omega$ almost every $n$. Then by Lemma 3.15, we know that $\mathrm{disp}_K(\rho_n(h^{*d} * f)x_n) < C^d(3 + \varepsilon_n + \frac{2+\varepsilon_n}{10})$ for $\omega$ almost every $n$. Since for $n$ large enough, $C^d(3 + \varepsilon_n + \frac{2+\varepsilon_n}{10}) < 4C^d$, this implies that, for $\omega$ almost every $n$, $\mathrm{disp}_K(\rho_n(h^{*d} * f)x_n) < 4C^d < C$, contradicting our assumptions. $\square$

**Remark:** In the discrete case, it is possible to prove Theorem 2.5 for the same function $h$ as in Proposition 3.1 and any $C > C_0$ from that proposition. The reader should note that this is no longer possible when $\Gamma$ is not discrete, but that $h$ can be replaced by $h^{*l}$ where $l$ is a constant depending only on $\Gamma, K$ and $h$.

3.4. **Proofs of Theorem 2.6.** We indicate the modifications to the proof of Theorem 2.5 needed to prove Theorem 2.6.

For detailed discussion and definitions about properties of Banach spaces, the reader should refer to [BL, Chapter 8] for positive definite functions and to [BL, Appendix A] for uniform convexity. We recall some consequences and definitions here. We will let $\mathcal{B}$ be a uniformly convex Banach space. Let $\mathcal{B}_1$, respectively $\mathcal{B}_1^*$, be the unit ball. Then there is a map $j : \mathcal{B}_1 \to \mathcal{B}_1^*$, called the *duality map* defined by letting $j(x)$ be the unique functional such that $\|j(x)\| = 1$ and $< j(x), x >= 1$ such that $j^{-1}$ is uniformly continuous. (For a proof of uniform continuity and estimates on the modulus of continuity in terms of the modulus of convexity see [BL, Appendix A].) An easy consequence of the definitions is the existence of a strictly increasing function $\zeta$ on $[0, 1]$ with $\zeta(0) = 0$ such that for any pair of vectors $v, w \in \mathcal{B}_1$ we have $< j(w), v > \leq 1 - \zeta(\varepsilon)$ if and only if $d(v, w) \geq \varepsilon$.

The following lemma is used in place of Lemma 3.2 above. For any representation $\sigma$ of $\Gamma$ on a Banach space $\mathcal{B}$, we denote by $\mathcal{B}^\sigma$ the set of $\sigma$ invariant vectors.

**Lemma 3.16.** *Let $\Gamma$ be a locally compact, compactly generated group and $K$ a compact generating set. Let $\mathcal{B}$ be a uniformly convex Banach space. Then for any unitary representation $\sigma$ of $\Gamma$ on $\mathcal{B}$ the following are equivalent:*

1. *there exists $M > 0$ such that for any $\delta \geq 0$, any $(K, \delta)$-almost invariant vector $v$ is within $M\delta$ of $\mathcal{B}^\sigma$;*
2. *for any function $f \in \mathcal{U}_2(\Gamma)$ there exists $0 < C < 1$ such that for any $v \in \mathcal{B}$ we have $d(\sigma(f)v, \mathcal{B}^\sigma) \leq C d(v, \mathcal{B}^\sigma)$.*

*Proof.* The proof that (2) implies (1) and the reverse implication in the discrete case are straightforward and similar to the proof of [M, Lemma III.1.1]. Therefore we only give an argument for (1) implies (2).

Fix a function $f \in \mathcal{U}_2(\Gamma)$, then there exists $\eta > 0$ such that $f(\gamma) > \eta$ for every $\gamma \in K^2$. Fix a vector $v$ with $d(v, \mathcal{B}^\sigma) > 0$. By re-scaling and changing basepoint, we can assume $d(v, \mathcal{B}^\sigma) = 1$ and in fact that $d(v, 0) = 1$ where 0 is the origin in $\mathcal{B}$. This uses the fact that $\mathcal{B}$ is uniformly convex which implies that there is a point in $\mathcal{B}^\sigma$ realizing $d(v, \mathcal{B}^\sigma)$. There exists $\gamma_0 \in K$ for which $d(\gamma\gamma_0 v, \gamma v) \geq \frac{1}{M}$. This implies that

$$\mu\{\gamma \in K^2 | d(\gamma v, w) \geq \frac{1}{2M}\} \geq \mu(K)$$

for any unit vector $w$. Applying $j(w)$ to $\sigma(f)v$ we have that

$$\int j(w)(f(\gamma)\sigma(\gamma)v) \leq 1 - \zeta(\frac{1}{2M})\mu(K)\eta.$$

Since this is true for any $w$, this implies that $d(\sigma(f)v,0)\leq 1-\zeta\frac{1}{2M}\mu(K)\eta$. ☐

We now state a replacement for Proposition 3.1.

**Proposition 3.17.** *If $\Gamma$ has property $(T)$ and $f\in\mathcal{U}_2(\Gamma)$ and $) < C_0 < 1$, there exist positive integers $M = M(f, C_0)$ and $m = m(f, C_0)$ such that, letting $h = f^{*m}$, for any generalized $L^p$ space $\mathcal{B}$, any continuous isometric action $\rho$ of $\Gamma$ on $\mathcal{B}$, and any $x\in\mathcal{B}$ we have*

(1) $d_{\mathcal{B}}(x, \rho(h)(x))\leq M \operatorname{disp}_K(x)$
(2) $\operatorname{disp}_K(\rho(h)(x))\leq C_0 \operatorname{disp}_K(x)$.

As this is essentially contained in [BFGM], we only provide a sketch.

*Sketch of Proof.* Let $g$ be the positive definite function on $\mathcal{B}$. Then $g$ defines maps $T_t : \mathcal{B}\to\mathcal{H}_1$ where $\mathcal{H}$ is a Hilbert space and $\mathcal{H}_1$ is the unit sphere. The map $T_t$ satisfies $< Tx, Ty >= g(t(x-y))$ where $g(x) = \exp(-\|x\|^p)$. One can then apply the standard proof that any affine action of a group with property $(T)$ on a Hilbert space has a fixed point, see for example [HV] or [De], to produce a $\Gamma$ fixed point on $\mathcal{B}$. In fact, the proof shows more. It produces a constant $C$, depending only on $\Gamma$ and $K$ such that the $\Gamma$ displacement of any point $y$ in $\mathcal{B}$ is bounded by $C$ times the $K$ displacement. This then implies that the distance from $y$ to a fixed point (the barycenter of $\Gamma\cdot y$) is bounded by a constant times the $K$ displacement of $y$. To verify these facts one uses the fact that $d_{\mathcal{H}}(T_t x, T_t y)^2 = 2td_{\mathcal{B}}(x,y)^p + O(t^2)$ for all $t$. The existence of $h$ then follows from Lemma 3.16 and an argument as in the proof of Proposition 3.4. ☐

The following fact about ultra-product spaces is left to the reader, compare [He].

**Proposition 3.18.** *Let $p_i$ be sequence of numbers with $1 < p_i < 2$ and $\mathcal{B}_i$ be a sequence of generalized $L^p$ spaces. Let $\omega$ be an ultra-filter and $p = \omega\text{-}\lim p_i$. Then the function $f(x) = \exp(-\|x\|^p)$ is positive definite on the ultra-product $\mathcal{B}$ of $\mathcal{B}_n$ and $\mathcal{B}$ is uniformly convex with the same modulus of convexity as $L^p$.*

*Proof of Theorem 2.6.* In the discrete case, the proof is a verbatim repetition of the proof of Theorem 2.5 with Proposition 3.1 replaced by Proposition 3.17 and Proposition 3.9 replace by Proposition 3.18.

For non-discrete $\Gamma$, one needs to verify that versions of Proposition 3.11 and 3.13 still hold, but modifying the proofs of those statements is straightforward if we assume all $\Gamma$ actions are affine. ☐

## 4. **Inner products on tensor spaces and existence of invariant metrics**

In this subsection we define intrinsic leafwise Sobolev structures on spaces of tensors on $T\mathfrak{F}$. For our applications, we need these structures on the space of functions and the space of symmetric two forms, but we develop it more generally. We define a family of norms on leafwise $C^r$ tensors and then complete with respect to the corresponding metric. For us the key fact about the norms we use is that they are invariant under isometries of the leafwise Riemannian metric. To illustrate the utility of this construction, we prove Theorem 1.3 using this construction and results from subsection 2.1 and 2.2.

As in subsection 2.3 we let $X$ be a locally compact, $\sigma$-compact, metric space, $\mathfrak{F}$ a foliation of $X$ by manifolds of dimension $n$, and $g_{\mathfrak{F}}$ a leafwise Riemannian metric. We also let $\nu_{\mathfrak{F}}$ denote the Riemannian volume (and corresponding measure) on leaves of $\mathfrak{F}$ and assume a transverse invariant measure $\nu$. We define norms on the set of tensors which are continuous globally and $C^r$ along leaves of $\mathfrak{F}$. The definitions given below are standard when $X$ is a single leaf. To make the norms intrinsic, we work with $k$-jets of sections of tensor bundles. The special case of functions, particularly important in our applications, is sections of the trivial one dimensional vector bundle which corresponds to tensors of the form $\otimes^0 T\mathfrak{F}$. Here $T\mathfrak{F}$ is the tangent bundle to the foliation $\mathfrak{F}$. We will denote by $\xi$ an arbitrary bundle of tensors in $T\mathfrak{F}$ and $\mathrm{Sect}^k(\xi)$ the space of sections of $\xi$ that are globally continuous and $C^k$ along leaves of $\mathfrak{F}$. By *globally continuous (resp. measurable) and $C^k$ along leaves of* $\mathfrak{F}$ we will always mean that an object is $C^k$ along leaves and varies continuously (resp. measurably) in the $C^k$ topology transverse to leaves. Particular examples include vector fields $\xi = T\mathfrak{F}$, symmetric two tensors $\xi = S^2(T\mathfrak{F}^*)$ or functions $\xi = X \times \mathbb{R}$.

**Remark:** For our first proof of Theorem 1.1, we allow one additional choice for $\xi$. Let $\xi = X \times \mathbb{R}^n$ be a trivial bundle. Given an action $\rho$ of $\Gamma$ on $X$, we normally would associate the trivial action on $\xi$. Instead we allow the possibility of the existence of finite dimensional unitary representation $\sigma$ of $\Gamma$ on $\mathbb{R}^n$, and define the action of $\rho$ on $\xi$ by $\rho(\gamma)(x, v) = (\rho(\gamma)x, \sigma(\gamma)v)$. Similarly for any perturbation $\rho'$ of $\rho$, we define the action $\rho'$ on $\xi$ by $\rho'(\gamma)(x, v) = (\rho'(\gamma)x, \sigma(\gamma)v)$.

Given $g_{\mathfrak{F}}$, there is a canonical Levi-Civita connection on $T\mathfrak{F}$ which we denote by $\nabla^T$ associated to the metric. For any choice of $\xi$, this defines a connection on $\xi$, see for example III.2 of [KN], which we will view as $\nabla^{\xi} : \mathrm{Sect}(T\mathfrak{F}) \times \mathrm{Sect}(\xi) \to \mathrm{Sect}(\xi)$. Note that $\nabla$ is always invariant under isometries of $g_{\mathfrak{F}}$. There is also a natural metric on $\xi$

associated to the metric on $T\mathfrak{F}$ (see for example section 20.8.3 in [D]). In the particular case of functions, the metric is any metric given by identifying all fibers with $\mathbb{R}$, the connection is given by $\nabla_X f = Xf$, and invariance of the connection is immediate.

We will let $J^k(\xi)$ denote the bundle of leafwise $k$-jets of sections of $\xi$. This is a bundle where the fiber over a point $x$ is the set of equivalence classes of continuous, leafwise $C^k$ sections where two sections are equivalent if they agree to order $k$ at the point $x$. We will denote the fiber by $J^k(\xi)_x$. There is a natural identification:

$$J^k(\xi)\simeq\bigoplus_{j=0}^{k}(S^j(T\mathfrak{F}^*)\otimes\xi).$$

As a special case of the discussion above, a metric on $T(\mathfrak{F})$ defines one on $S^j(T\mathfrak{F}^*)$ for all $j$ and together with the metric on $\xi$, this identification induces a metric on $J^k(\xi)$. We briefly review the identification above to show that this metric is indeed invariant under isometries of $g_{\mathfrak{F}}$. The exposition that follows draws mostly from section 9 of [P1], where the interested reader can find more proofs and explicit constructions.

We can view the leafwise Levi-Civita connection on $T\mathfrak{F}$ as a map $\nabla^T : \mathrm{Sect}(T\mathfrak{F})\to\mathrm{Sect}(T\mathfrak{F}^*)\otimes\mathrm{Sect}(T\mathfrak{F})$. By the discussion above, we also have a connection $\nabla^{T*} : \mathrm{Sect}(T\mathfrak{F}^*)\to\mathrm{Sect}(\otimes^2 T\mathfrak{F}^*)$. Similarly the connection on $\xi$ can be viewed as a map $\nabla^\xi : \mathrm{Sect}(\xi)\to\mathrm{Sect}(T\mathfrak{F}^*\otimes\xi)$ by viewing $\nabla_X\sigma$ as a one form with $X$ as the variable.

Similarly, we can define a canonical covariant derivative

$$\nabla^{(i)} : \mathrm{Sect}(\otimes^i(T\mathfrak{F}^*)\otimes\xi)\to\mathrm{Sect}(\otimes^{i+1}(T\mathfrak{F}^*)\otimes\xi)$$

via the formula

$$\nabla^{(i)}(V_1\otimes\cdots\otimes V_i\otimes f) = \sum_{j=1}^{i}(V_1\otimes\cdots\otimes\nabla^{T*}V_j\otimes\cdots\otimes V_i\otimes f)+V_1\otimes\cdots\otimes V_i\otimes\nabla^\xi f$$

where $f$ is a section of $\xi$ and the $V_i$ are elements of $T\mathfrak{F}^*$. The composition $\nabla^{(k-1)}\cdots\nabla^{(1)}\nabla : \mathrm{Sect}(\xi)\to\mathrm{Sect}(\otimes^k(T\mathfrak{F}^*)\otimes\xi)$ is called a $k$th covariant derivative.

We now define the total covariant derivative. First let $S^{(k)}$ be the natural symmetrization operator from the $k$th tensor power of $T\mathfrak{F}$ to $S^k(T\mathfrak{F})$ the $k$th symmetric power. We define the $k$th total differential $D_k = S^{(k)}_*\nabla^{k-1}\cdots\nabla^{(1)}\nabla$. In other words, the $k$th total differential is the symmetrization of the $k$th covariant derivative. The isomorphism mentioned above $J^k(\xi)\simeq\bigoplus_{m=0}^{k}(S^j(T\mathfrak{F}^*)\otimes\xi)$ is given by the map $j^k(v) = \{D_m(v)\}_{0\leq m\leq k}$. We then define the metric on $J^k(\mathfrak{F})$ via this

isomorphism and abuse notation by calling it $g$. Define a family of norms on $\mathrm{Sect}(J^k(\xi))$ via:

$$\|u\|_p^p = \int_X g(u_x, u_x)^{\frac{p}{2}} d\nu_{\mathfrak{F}} d\nu,$$

for every $u \in \mathrm{Sect}(J^k(\xi))$. Since elements of $\mathrm{Sect}^k(\xi)$ define elements of $\mathrm{Sect}(J^k(\xi))$ we can restrict this to an inner product on $\mathrm{Sect}^k(\xi)$ defined by

$$\|u\|_{k,p}^p = \int_X < j^k(u), j^k(u) >^{\frac{p}{2}} d\nu_{\mathfrak{F}} d\nu,$$

for all $u \in \mathrm{Sect}^k(\xi)$.

**Notational convention:** Throughout this paper when $f$ is leafwise smooth homeomorphism of $(X, \mathfrak{F})$ and therefore induces a map on functions or sections of a tensor bundle $\xi$ over $X$, we abuse notations by writing $f$ for the map on functions or sections. This remark also applies to group actions.

**Proposition 4.1.** *Let $f$ be a leafwise isometry of $(X, \mathfrak{F}, g_{\mathfrak{F}})$ which preserves the transverse invariant measure $\nu$. Then the action of $f$ on* $\mathrm{Sect}^k(\xi)$ *and* $\mathrm{Sect}(J^k(\xi))$ *preserves all of the norms defined above.*

*Proof.* This is clear from the definition of the inner product and the fact that isometries of $g_{\mathfrak{F}}$ commute with all the differential operators used in the construction and that $f$ preserves the measure in the integral above. □

We now have norms defined $\mathrm{Sect}(J^k(\xi))$ which restrict to norms defined on $\mathrm{Sect}^k(\xi)$. We define distance functions on $\mathrm{Sect}(J^k(\xi))$ by $d_p(u,v) = \|u - v\|_p$ and refer to the completion with respect to this metric as $L^p(J^k(\xi))$. Note that $d_p$ restricts to a metric $d_{p,k}$ on $\mathrm{Sect}^k(\xi)$ that is exactly the metric induced by $\|\cdot\|_{k,p}$. Completing $\mathrm{Sect}^k(\xi)$ with respect to $d_{p,k}$ we obtain a standard Sobolev completion of that space, a Banach subspace of $L^p(J^k(\xi))$, which we denote by $L^{p,k}(\xi, \mathfrak{F})$. If the foliation is the trivial foliation by a single leaf $X$, we omit the $\mathfrak{F}$ and simply write $L^{p,k}(\xi)$ and $L^p(J^k(\xi))$. In the special case of functions, we use the notation $L^{p,k}(X, \mathfrak{F}), L^p(J^k(X))$ or $L^{p,k}(X)$ in place of $L^{p,k}(X \times \mathbb{R}, \mathfrak{F}), L^p(J^k(X \times \mathbb{R})$ or $L^{p,k}(X \times \mathbb{R})$ respectively. It is clear that if $f$ is a homeomorphism of $X$ as in Proposition 4.1 then the action of $f$ on $\mathrm{Sect}(J^k(\xi))$ and $\mathrm{Sect}^k(\xi)$ extend to isometric actions on $L^p(\mathrm{Sect}(J^k(\xi))$ and $L^{p,k}(\xi)$. Since any $u \in L^{p,k}(\mathfrak{F}, \xi)$ is a limit of $u_i \in \mathrm{Sect}^k(\xi)$ with respect to the norm above, it follows that $j^k(u_i)$ converge in $L^p(J^k(\xi))$ to a section we denote by $j^k(u)$ and call the *weak $k$-jet* of $u$.

We also have the following fact about perturbations of isometric actions which will be used heavily in the next section.

**Proposition 4.2.** *Let $f$ be a leafwise isometry of $(X, \mathfrak{F}, g_{\mathfrak{F}})$ which preserves $\nu$ and let $s \in \mathrm{Sect}^k(\xi)$ be $f$ invariant. If $\xi$ is a trivial bundle, we let $l = k$, if $\xi$ is non-trivial, we let $l = k+1$. For any $p_0 > 1, \varepsilon > 0$ and $\delta > 0$ there exists a neighborhood $U$ of the identity in $\mathrm{Diff}^l_\nu(X, \mathfrak{F})$ such that if $f'$ is an $(U, C^l)$-foliated perturbation of $f$ then*

1. *for any $p \leq p_0$, the action of $f'$ on $L^p(J^k(\xi))$ (and therefore $L^{p,k}(\xi, \mathfrak{F})$) is by $\varepsilon$-almost isometries,*
2. *the $f'$ displacement of $s$ in $L^{p,k}(\xi, \mathfrak{F})$ is less than $\delta$,*
3. *if $V \subset X$ is any $f'$ invariant set of positive measure, then the action of $f'$ on $L^p(J^k(\xi))|_V$ (and therefore $L^{p,k}(\mathfrak{F}, \xi)|_V$) is by $\varepsilon$-almost isometries,*
4. *if $V \subset X$ is any $f'$ invariant set of positive measure, then the $f'$ displacement of $s|_V$ in $L^{p,k}(\mathfrak{F}, \xi)|_V$ is less than $\delta\mu(V)$.*

*Furthermore if $H$ is a topological group and $\rho$ is a continuous leafwise isometric action of $H$ on $X$, then the resulting $H$ action on $L^p(J^k(\xi))$ (and therefore $L^{p,k}(\xi, \mathfrak{F})$) is continuous. The same is true for any continuous $(U, C^k)$-foliated perturbation $\rho'$ of $\rho$*

**Remarks:**

1. The choice of $l$ is required since while $C^l$ diffeomorphisms act on $C^l$ functions on $X$, they only act on $C^{l-1}$ sections of any non-trivial tensor bundle $\xi$. It is easy to verify that if $f$ is a $C^l$ diffeomorphism and $s$ is a $C^{l-1}$ section of $\xi$, then $j^{l-1}(s \circ f) = j^{l-1}(s) \circ j^l(f)$.
2. Since if $s$ is in $\mathrm{Sect}^\infty(\xi)$ then $s$ is in $\mathrm{Sect}^k(\xi)$ we can use Proposition 4.2 to study translates of $C^\infty$ sections inside Sobolev spaces.
3. There is no better statement for the $C^\infty$ case, since a $C^\infty$ neighborhood of $f$ is exactly a $C^n$ neighborhood of $f$ for some large integer $n$. If $f'$ is close to $f$ in the $C^n$ topology but not the $C^{n+1}$ topology, then even if $f'$ is $C^\infty$, $f'$ is not be an $\varepsilon$-almost isometry on any space defined using more than $n$ derivatives. Therefore we can only obtain an estimates for the $f'$ action on spaces whose norms depend on no more than $n$ derivatives.

*Proof.* First given $\varepsilon$, we find $U$ such that $(1-\varepsilon)\|s\|_p \leq \|s \circ f'\|_p \leq (1-\varepsilon)\|s\|_p$ for any $s \in L^p(\mathrm{Sect}(\xi))$ and for any $f'$ which is $(U, C^l)$-close to $f$. Since continuous sections are dense in $L^p(\mathrm{Sect}(\xi))$ (this follows from the fact that continuous functions are dense in $L^p$), we can assume that $s$ is continuous. We can write $s \circ f'$ as $s \circ (f \circ f^{-1}) \circ f' =$

$(s{\circ}f){\circ}(f^{-1}{\circ}f')$. Since $f$ is an isometry of $L^p(\mathrm{Sect}(\xi))$ it suffices to show that $(1-\varepsilon)\|s\|_p \leq \|s{\circ}(f^{-1}{\circ}f'){\cdot}\|_p \leq (1-\varepsilon)\|s\|_p$ for leafwise smooth $s$. For any $\eta > 0$, we can choose $U$, an open set in $\mathrm{Diff}^l(X, \mathfrak{F})$ containing the identity, such that $1-\eta < \|j^l(f^{-1}{\circ}f')(x)\| < 1+\eta$ for all $x$. Then the chain rule implies the pointwise bound

$$(1-\eta)\|j^k(s)(x)\| \leq \|j^k(s{\circ}(f^{-1}{\circ}f'))(f^{-1}(f'(x))\| \leq (1+\eta)\|j^k(s)(x)\|.$$

We further restrict $U$ so that the Jacobian of $f'$ along $\mathfrak{F}$ is bounded between $1+\eta$ and $1-\eta$, and then the result follows from the fact that $f'$ preserves the transverse measure $\nu$ provided $\varepsilon < (1+\eta)^{p+1} - 1$. This argument also verifies that $f'$ acts by $\varepsilon$-almost isometries on $L^p(\mathrm{Sect}(\xi))|_V$ for any $V \subset X$ of positive measure.

The remaining conclusions follow from the fact that $\mathrm{Diff}^k(X, \mathfrak{F})$ acts continuously on $\mathrm{Sect}(J^k(\xi))$ and $\mathrm{Sect}^k(\xi)$ and therefore on $L^p(J^k(\xi))$ and $L^{p,k}(\xi)$. □

In order to obtain optimal results, we need to make precise some notions of Hölder regularity in order to have a norm on $\mathrm{Sect}^k(\xi)$ where $k$ is not integral. For the remainder of this section, we allow the possibility that $k$ is not integral and let $k'$ to denote the greatest integer less than $k$. Given $x{\in}X, y{\in}\mathfrak{L}_x$ and a piecewise $C^1$ curve $c$ in $\mathfrak{L}_x$ joining $x$ to $y$, for any natural vector bundle $V$ over $X$, we denote the parallel translation of $v{\in}V_y$ to $V_x$ by $P^x_y v$ and by $l(c)$ the length of $c$. We then define

$$\|s\|_k = \|s\|_{k'} + \sup \frac{\|P^x_y j^{k'}(s)(y) - j^{k'}(s)(x)\|}{l(c)^{k-k'}}$$

where $k'$ is the least integer not greater than $k$ and the supremum is taken over $x{\in}X$, $y{\in}\mathfrak{L}_x$ and piecewise $C^1$ curves $c$ in $\mathfrak{L}_x$ joining $x$ to $y$. It is easy to verify that this definition agrees with the usual one in the Euclidean case.

We now also make precise the $C^k$ size of a $C^k$ map $f : Z{\to}Z$ where $Z$ is a Riemannian manifold and $k$ is not an integer. This notion is needed to make precise the conclusion (4) of Theorem 2.11. We already have a notion of pointwise $C^{k'}$ size, defined in subsection 7.2, which we denote by $\|j^{k'}(f)(x)\|$. Recall that $j^{k'}(f)(x) : J^{k'}(Z, \mathbb{R})_x{\to}J^{k'}(Z, \mathbb{R})_{f(x)}$ is a linear map between vector spaces. Given a curve $c$ in $Z$, we can compose $j^{k'}(f)(x)$ with parallel translation $P^x_{f(x)}$ along $c$ to obtain a self-map $P^x_{f(x)}{\circ}j^{k'}(f)(x)$ of $J^{k'}(Z, \mathbb{R})_x$. We define the $C^k$ size of $f$ to be

$$\|f\|_k = \sup_x \|j^{k'}(f)(x)\| + \sup \frac{\|P^x_{f(x)}{\circ}j^{k'}(f)(x)\|}{l(c)^{k-k'}}$$

where the supremum is taken over $x \in X$, $y \in Z$ and piecewise $C^1$ curves $c$ in $Z$ joining $x$ to $y$. We can also measure the $C^{k'}$ size of $f$ on any subset $U$ of $Z$ by restricting the above supremum to $x \in U$.

**Proposition 4.3.** *Let $(X, \mathfrak{F})$ be a compact foliated space and $g_{\mathfrak{F}}$ a continuous, leafwise smooth metric on $(X, \mathfrak{F})$. Then for any $\xi$ and any $p > 1$, there are uniformly bounded inclusions $L^{p,k}(\tilde{\mathfrak{L}}_x, \xi) \subset \mathrm{Sect}^{k-\frac{d}{p}}(\xi|_{\tilde{\mathfrak{L}}_x})$ for all $x$, where $d = \dim(Z)$ and $\tilde{\mathfrak{L}}_x$ is any covering space of the leaf through $x$.*

*Proof.* The standard Sobolev embedding theorems provide an bounded inclusion of $L^{p,k}(\mathbb{R}^d)$ in $C^{k-\frac{d}{p}}(\mathbb{R}^d)$ which easily implies a bounded embedding of $L^{p,k}(\mathbb{R}^d, \mathbb{R}^n)$ in $C^{k-\frac{d}{p}}(\mathbb{R}^d, \mathbb{R}^n)$. Compactness of $X$ and the fact that $g_{\mathfrak{F}}$ is continuous and leafwise smooth, imply that we can cover $X$ with a finite collection of charts $(U_i, \phi_i)$ with $\phi_i(U_i) = V_i \times B(0, c)$ and such that there is a uniform bound on the resulting inclusions

$$L^{p,k}(\mathfrak{F}, \xi|_{B_{\mathfrak{F}}(v_i,c)}) \subset L^{p,k}(B(0,c), \mathbb{R}^n)$$

and

$$C^{k-\frac{p}{d}}(B(0,c), \mathbb{R}^n) \subset \mathrm{Sect}^{k-\frac{d}{p}}(\xi|_{B_{\mathfrak{F}}(v_i,c)})$$

for every $U_i$ and every $v_i \in V_i$, where $B_{\mathfrak{F}}(v_i, c) = \phi^{-1}(v_i \times B(0, c))$. So we have uniformly bounded embeddings

$$L^{p,k}(\mathfrak{F}, \xi|_{B_{\mathfrak{F}}(v_i,c)}) \subset L^{p,k}(B(0,c), \mathbb{R}^n) \subset C^{k-\frac{d}{p}}(B(0,c), \mathbb{R}^n) \subset \mathrm{Sect}^{k-\frac{d}{p}}(\xi|_{B_{\mathfrak{F}}(v_i,c)})$$

for every $U_i$ and every $v_i \in V_i$ which suffices to complete the proof. It is easy to see that the same bound holds for any cover $\tilde{\mathfrak{L}}_x \to \mathfrak{L}_x$. If $k - \frac{d}{p}$ is not integral, this does not immediately yield the desired result, since we only have a Hölder bound at small scales. However, since we have a global bound on the $C^0$ norm, it is easy to convert this small scale Hölder bound to a worse Hölder bound on all scales. More precise estimates can be obtained by following the standard proofs of Hölder regularity in the Sobolev embedding theorems. □

If we are studying perturbations $\rho'$ of an action $\rho$, in order to obtain optimal regularity in all proofs, we will need to know that a certain section $s'$ invariant under $\rho'$ is close in $L^{p,k}$ type Sobolev spaces to certain $\rho$ invariant section $s$. The difficulty here is to show that $s'$ is both invariant under $\rho'$ and close in $L^{p,k}$ for $p > 2$ simultaneously. To show this, we will require the following elementary fact.

**Lemma 4.4.** *Let $(X, \mu)$ be a measure space and $V$ a finite dimensional vector space. Assume that $f_n \in L^p(X, \mu, V)$ converge in $L^p$ to a function*

*$f$. Further assume that $\|f_n - f_{n+1}\|_p \leq C^n$ where $0 < C < 1$. Then $f_n$ converges pointwise almost everywhere to $f$.*

*Proof.* Let $X_n = \{x \| f_n(x) - f_{n+1}(x)| > C^{n/2p}\}$, since $\|f_n - f_{n+1}\|_p \leq C^n$, it follows that $\mu(X_n) < C^{n/2}$. Then $\{f_n\}$ converges pointwise on the complement of $X_\infty = \cap_{n=1}^\infty \cup_{k=n}^\infty X_n$. The lemma follows from the Borel-Cantelli lemma, since $\sum_n \mu(X_n)$ converges, so $\mu(X_\infty) = 0$. □

Many of our uses of this fact could, with slight rewording, be deduced from the fact that if a sequence of functions $\{f_n\}$ converges to a function $f^p$ $L^p$ and converges to a function $f^q$ in $L^q$ then $f^p = f^q$ almost everywhere. However, the full strength of Lemma 4.4 is required in the proof of Theorem 2.11.

**Lemma 4.5.** *Let $\Gamma$ be a locally compact, $\sigma$-compact group with property $(T)$ generated by a compact set $K$, and let $\rho$ be a leafwise isometric action of $\Gamma$ on $(X, \mathfrak{F}, g_{\mathfrak{F}})$ and $s$ a $\rho$ invariant section in $\mathrm{Sect}^k(\xi)$. Let $l$ be as in Proposition 4.2. For any $p \geq 2, \eta > 0, F > 0$ and $0 < C < 1$, there exists a neighborhood $U$ of the identity in $\mathrm{Diff}^l_\nu(X, \mathfrak{F})$ and a function $h = h(p)$ in $\mathcal{U}(\Gamma)$ such that if $\rho'$ is a $(U, C^l)$-foliated perturbation of $\rho$,*

1. *$\rho'(h)^n s$ converge pointwise almost everywhere to a $\rho'$ invariant section $s'$ in $L^{p,k}(\xi)$,*
2. *$\|\rho'(h)^n s - s'\|_{p,k} \leq \eta$ for all $n \geq 0$ and,*
3. *$\|\rho(h)^{n+1} s - \rho(h)^n s\|_{p,k} < C^n F$ for all $n \geq 0$.*

**Remarks:**

1. For many applications we only need conclusion (1) and the case of (2) where $n = 0$, i.e. that $\|s' - s\|_{p,k} \leq \eta$.
2. The reason we do not obtain these estimates in $L^{p,k}$ for all $p, k$ when $\rho'$ is $C^\infty$ close to $\rho$ is explained following Proposition 4.2.

*Proof.* Given $p \geq 2, \varepsilon > 0$ and $\delta = \min(\frac{F}{M}, \frac{\eta(1-C)}{MC})$ for $M$ to be specified below, by Proposition 4.2 there is a neighborhood $U$ in $\mathrm{Diff}^l(X)$ such that for any $(U, C^l)$-perturbation $\rho'$ of $\rho$, it follows that $\rho'(k)$ is an $\varepsilon$-almost isometry of $L^p(\mathrm{Sect}^k(\xi, \mathfrak{F}))$ (and therefore of $L^{p,k}(\xi, \mathfrak{F})$) for any $p < p_0$ and the $\mathrm{disp}_K(s) < \delta'$ in any of these spaces.

We choose $h \in \mathcal{U}(\Gamma)$ satisfying both Theorem 2.1 and Corollary 2.8. Then Theorem 2.1 shows that $\rho'(h)^n s$ converges exponentially to $s'$ in $L^{2,k}(\xi, \mathfrak{F})$, which by Lemma 4.4, implies that $\rho'(h)^n s$ converges pointwise almost everywhere to $s'$. Then applying Corollary 2.8 there is a constant $M = M(h, C, p)$, such that $\|\rho(h)^{n+1} s - \rho(h)^n s\| < C^n M \delta$ and therefore $s' \in L^{p,k}(\xi, \mathfrak{F})$ and $\|\rho(h)^n s - s'\|_{p,k} \leq \frac{MC}{1-C} \delta$ for all $n \geq 0$. By our choice of $\delta$ we have $\|\rho(h)^n s - s'\| \leq \eta$ and $\|\rho(h)^{n+1} s - \rho(h)^n s\| < C^n F$ for all $n \geq 0$ as desired. □

To illustrate the application of the results in section 2.1 to perturbations of isometric and leafwise isometric actions, we now prove Theorem 1.3 from the introduction.

*Proof of Theorem 1.3.* We have an action $\rho$ of a group $\Gamma$ with property $(T)$ on compact manifold $X$ preserving a Riemannian metric $g$. We view $g$ as a section of the (positive cone in) the bundle of symmetric two tensors $S^2(TX)$. Fix a generating set $K$ of $\Gamma$ and a choice of $\eta > 0$ to be specified below. Given $\varepsilon > 0$ satisfying the hypotheses of Theorem 1.6 and $\delta > 0$ to be specified below, by Proposition 4.2 there is a neighborhood $U$ in $\mathrm{Diff}^{k+1}(X)$ such that for any $(U, C^{k+1})$-perturbation $\rho'$ of $\rho$, it follows that $\rho'(k)$ is an $\varepsilon$-almost isometry of $L^{2,k}(S^2(TX))$ and the $K$ displacement of $g$ is less than $\delta$ in this space. Theorem 1.6 then implies that there is a number $C > 0$ depending only on $\Gamma$ and $K$ and a $\rho'(\Gamma)$ invariant section $g' \in L^{2,k}(S^2(TX))$ with $\|g - g'\|_{2,k} \leq \eta$ where $\eta = C\delta$ is specified below. To obtain optimal regularity, we choose $p > \frac{d}{\kappa}$ and let $U$ satisfy Lemma 4.5 for $p$ and $\eta$ specified below and then Lemma 4.5 implies that there is a $\rho'$ invariant section $g'$ such that $\|g - g'\|_{p,k} < \eta$.

By Proposition 4.3, this implies that $\|g - g'\|_{k-\frac{d}{p}} \leq C'\eta$ where $d = \dim(X)$ and $C'$ depends only on $X$ and $g$. Since the cone of positive definite metrics is open in $S^2(TX)$, we can choose $\eta$ depending only on $p$ and $g$, so $g'$ is a $C^{k-\frac{d}{p}}$ Riemannian metric on $X$, invariant under $\rho'(\Gamma)$ and $C^{k-\frac{d}{p}}$ close to $g$. □

## 5. **Property $(T)$ and conjugacy**

In this section we prove Theorem 1.1 using Theorem 1.6 and Lemma 4.5. In this section, we only consider $C^k$ perturbations. The additional arguments required for the $C^\infty$ case are in section 6.

### 5.1. **A proof of Theorem 1.1.**

We begin by noting a classical fact about isometric actions.

**Proposition 5.1.** *Let $0 \leq k \leq \infty$, let $X$ be a compact $C^k$ manifold and $\rho$ a $C^k$ action of $\Gamma$ on $X$ such that the image of $\Gamma$ in $\mathrm{Diff}^k(X)$ is pre-compact. Then there is a positive integer $n$, a homomorphism $\sigma : \Gamma \to O(n)$ and a $\Gamma$ equivariant $C^k$ embedding $s : X \to \mathbb{R}^n$.*

**Remark:** For our applications, the fact that $\Gamma$ is precompact in $\mathrm{Diff}^k(X)$ follows from the fact that the isometry group of a compact Riemannian manifold is compact.

*Proof.* Let $C$ be the closure of $\Gamma$ in $\mathrm{Diff}^k(X)$. For $C$ this is the Mostow-Palais theorem [Mo, P2]. More precisely, for all $k$, Mostow has proven that, for some $n$, there is a map $C \to O(n)$ and a $C^k$ equivariant embedding of $X$ into the Euclidean space $\mathbb{R}^n$. For $k=0$ this is the main result of [Mo], for $k>0$, it is proven in section 7.4 of that paper by a different method. For $k=\infty$, the same result is proven in [P2] using the fact that $C$ preserves a $C^\infty$ Riemannian metric. (Mostow's proofs do not explicitly use the existence of an invariant metric. In the $C^k$ case, Palais' method produces an equivariant embedding of lower regularity.) □

**Remark:** If $k \geq 2$, and $\Gamma$ preserves a $C^{k,\alpha}$ Riemannian metric $g$, then one can prove the $C^k$ version of the above theorem by approximating any embedding of $\Gamma$ in $\mathbb{R}^n$ with an embedding defined by eigenfunctions of the Laplacian.

Given $\sigma$ and $n$ as in Proposition 5.1, we define a trivial bundle $\xi = X \times \mathbb{R}^n$ with $\Gamma$ action $\rho(\gamma)(x,v) = (\rho(\gamma)x, \sigma(\gamma)v)$. The conclusion of Proposition 5.1 is then equivalent to the existence of a $\Gamma$ invariant section $s : X \to \xi$. We will show that the perturbed action preserves a section $s'$ close to $s$ and then use the following lemma to produce the conjugacy. Given a compact manifold $Y \subset \mathbb{R}^n$, there is a neighborhood $U$ of $Y$ in the normal bundle of $Y$ in $\mathbb{R}^n$ such that the exponential map $\exp : U \to \mathbb{R}^n$ defined by $\exp(x,v) = x+v$ is a diffeomorphism. The closest point projection $\phi$ from $\exp(U)$ to $Y$ is then $C^\infty$ (resp. $C^{n-1}$) when $Y$ is $C^\infty$ (resp. $C^n$). This yields the following:

**Lemma 5.2.** *Let $s : X \to \mathbb{R}^n$ be a $C^\infty$ embedding, then there exists $\eta$ such that for any integers $l \geq k \geq 1$ and any $s' : X \to \mathbb{R}^n$ a $C^l$ map with $\|s'-s\|_k \leq \eta$, the map $\psi = s^{-1} \circ \phi \circ s'$ is a $C^k$ small, $C^l$ diffeomorphism of $X$. Furthermore as $\eta \to 0$, the map $\psi$ tends to the identity map in the $C^k$ topology.*

We now prove Theorem 1.1. The reader who is only interested in a result, and not a result with optimal regularity, may ignore the last sentence of each paragraph and read the second paragraph assuming $p=2$. For any perturbation $\rho'$ of $\rho$, we define an action $\rho'$ on $\xi$ by $\rho'(\gamma)(x,v) = (\rho'(\gamma)x, \sigma(\gamma)v)$.

*Proof of Theorem 1.1.* Fix a generating set $K$ for $\Gamma$ and a constant $\eta > 0$ to be specified below. By Proposition 4.2, given $\varepsilon > 0$ satisfying the hypotheses of Theorem 1.6 and $\delta > 0$ specified below, we can choose a neighborhood of the identity $U \subset \mathrm{Diff}^k(X)$ such that for any $\rho'$ that is $(U, C^k)$ close to $\rho$, the map $\rho'(\gamma)$ is a $\varepsilon$-almost isometry of $L^{2,k}(\xi)$ for any $\gamma \in K$ and such that $\mathrm{disp}_K(s) < \delta$ in $L^{2,k}(\xi)$. Theorem

1.6 then implies that there is a $\rho'(\Gamma)$ invariant section $s'\in L^{2,k}(\xi)$ with $\|s-s'\|_{2,k}\leq\eta$ where $\eta = C\delta$ and $C > 0$ depends only on $\Gamma$ and $K$. To obtain a $C^{k-\kappa}$ conjugacy, we choose $p < \frac{d}{\kappa}$ and choose $U$ to satisfy Lemma 4.5 for our choices of $\eta$ and $p$ and then Lemma 4.5 implies that there is a $\rho'$ invariant section $s'$ with $\|s-s'\|_{p,k} < \eta$.

Proposition 4.3 then implies that $s'$ is $C^{k-\frac{d}{p}}$ and that $\|s-s'\|_{k-\frac{d}{p}} < C_0\eta$ where $C_0$ is an absolute constant depending only on $X, p$ and $g$. We can view $s' : X\to\mathbb{R}^n$ as a $\Gamma$ equivariant $C^{k-\frac{d}{p}}$ map from $X$ to $\mathbb{R}^n$ where the action on $X$ is given by $\rho'$ and the action on $\mathbb{R}^n$ is given by $\sigma$. By choosing $\eta$ (and therefore $U$) sufficiently small, the map $s' : X\to\mathbb{R}^n$ is $C^{k-\frac{d}{p}}$ close to $s$. Then by Lemma 5.2, the map $\psi = s^{-1}\circ\phi\circ s'$ is a $C^{k-\frac{d}{p}}$ small diffeomorphism of $X$. Since $s, s'$ and $\phi$ are all equivariant, $\psi$ is a conjugacy between the $\rho'(\Gamma)$ action on $X$ and the $\rho(\Gamma)$ action on $X$. □

5.2. **Another Proof of Theorem 1.1.** In this section we give another proof of Theorem 1.1 which gives somewhat lower regularity, but which generalizes to prove Theorem 2.11.

We will denote points in $X\times X$ by $(x_1, x_2)$ and denote the diagonal in $X\times X$ by $\Delta(X)$. Given any group $\Gamma$ acting on a manifold $X$, we will denote by $\bar{\rho}$ the diagonal action of $\Gamma$ on $X\times X$ given by $\bar{\rho}(\gamma)(x_1, x_2) = (\rho(\gamma)x_1, \rho(\gamma)x_2)$.

We begin with two elementary facts. Recall that a normal neighborhood of $x$ is the image under the Riemannian exponential map of an open ball $B$ in $T_xX$, such that $\exp_x|_B$ is a diffeomorphism and $d_X(x, \exp(v)) = d_{T_xX}(0, v)$. It is immediate that $d(x,\cdot)^2$ is a smooth function on any normal neighborhood of $x$. Let $N(x)$ be the maximal radius of a normal neighborhood of $x$ in $X$. On $X\times X$ we have a Riemannian metric on $g\times g$ and the induced distance function. By $B(x,\varepsilon)$ we denote the ball of radius $\varepsilon$ around $x$ in $X$, by $B((x,y),\varepsilon)$ the ball of radius $\varepsilon$ around $(x,y)$ in $X\times X$. Since $\{x\}\times X$ is totally geodesic in $X\times X$, we have $B((x,x),\varepsilon)\cap\{x\}\times X = B(x,\varepsilon)$.

**Proposition 5.3.** *Let $\rho$ be an isometric action of any group $\Gamma$ on any Riemannian manifold $X$. If we further assume the function $N(x) > d$ for some $d > 0$ and all $x\in X$, then for any $0 < \varepsilon < \frac{d}{2}$, there exists an invariant smooth function $f$ on $X\times X$ such that:*

1. *$f$ takes the value $0$ on $\Delta(X)$,*
2. *$f\geq 0$ and $f(x,y) > 0$ if $x\neq y$,*
3. *for any $x\in X$, the restriction of $f$ to $\{x\}\times X$ satisfies $f\geq 1$ outside of $B((x,x),\varepsilon)$,*

(4) *the Hessian of $f$ restricted to $\{x\}\times X$ is positive definite on the closure of $B((x,x),\varepsilon)$*

*Proof.* The action $\bar{\rho}$ leaves invariant any function on $X\times X$ which is a function of $d(x_1,x_2)$ and we define $f$ as such a function. To define $f$ we first define $f_{x_0}:X\to\mathbb{R}$ by $\frac{1}{\varepsilon^2}d(x,x_0)^2$. Our assumptions on $\varepsilon$ imply that $B(x_0,2\varepsilon)$ is contained in a normal neighborhood of $x_0$ and so $d(x,x_0)^2$ is a smooth function of $x$ and $x_0$ inside $B(x_0,2\varepsilon)$, see for example [KN, IV.3.6]. It is clear that $f(x_1,x_2)=f_{x_1}(x_2)$ satisfies all the requirements except smoothness on points at distance greater than $2\varepsilon$ from $\Delta(X)$. We merely need to change $f_{x_0}$ outside $B(x_0,\varepsilon)$ to produce a smooth $f_{x_0}$ while keeping $f_0\geq 1$ outside $B(x_0,\varepsilon)$. This is easily done by choosing any smooth function $g:\mathbb{R}\to\mathbb{R}$ such that $g$ agrees with $\frac{1}{\varepsilon^2}x^2$ to all orders for all $x\leq\varepsilon$ with $g\geq 1$ for all $x>\varepsilon$ and $g=1$ for all $x\geq 2\varepsilon$. We then let $f(x_1,x_2)=g(d(x_1,x_2))$. □

**Proposition 5.4.** *Let $X$ be a Riemannian manifold and $f$ a function on $X\times X$ such that:*

(1) *$f$ takes the value $0$ on $\Delta(X)$,*
(2) *$f\geq 0$,*
(3) *for any $x\in X$, the restriction of $f$ to $\{x\}\times X$ satisfies $f\geq 1$ outside $B((x,x),\varepsilon)$,*
(4) *the Hessian of $f$ restricted to $\{x\}\times X$ is positive definite on the closure of $B((x,x),\varepsilon)$.*

*Let $f'$ be a function which is $C^k$ close to $f$ where $k\geq 2$. Then for every $x$, the restriction of $f'$ to $\{x\}\times X$ has a unique global minimum at a point $(x,x')$ which is close to the point $(x,x)$. Furthermore, if we let $X'=\{(x,x')|f'(x')$ is the global minimum of $f'$ on $\{x\}\times X\}$ then $X'$ is a $C^{k-1}$ embedded copy of $X$ which is $C^{k-1}$ close to $\Delta(X)$.*

**Remark:** The last statement of the Proposition means that $X'$ is close to $\Delta(X)$ in the $C^{k-1}$ topology on $C^{k-1}$ submanifolds of $X\times X$. This actually suffices to imply that $X'$ is diffeomorphic to $X$ by a normal projection argument like the one used to prove Lemma 5.2.

*Proof.* Let $B=\{x\}\times B(x,\varepsilon)\subset\{x\}\times X$. We first verify the existence of $(x,x')$ in $B$. Since $f'$ is $C^k$ close to $f$ we have that $f'\geq\frac{1}{2}$ outside $B$ and $f'$ is close to zero near $(x,x)$. We look at all local minima of $f'$ on $\bar{B}$, the closure of $B$. Since $f'$ is close to $f$, at least one such minimum occurs in $B$. Since $k\geq 2$, if $f'$ is sufficiently $C^k$ close to $f$, the Hessian of $f'$ is positive definite on $B$, which implies there is exactly one local minimum on $B$, say at $(x,x')$. Since $f'$ is $C^k$ close to $f$, it is easy to see that $f'(x,x')$ must be close to zero and that $f'(x,y)$ must be close

to one if $f'|_{\{x\}\times X}$ has a local minimum at $(x, y)$ and $y$ is not in $\bar{B}$. Therefore $x'$ is the unique global minimum of $f'$ on $\{x\}\times X$.

Given a function $g : X\times X\to\mathbb{R}$ we denote by $D_2g$ the derivative with respect to the second variable, which is naturally a map from $X\times TX$ to $\mathbb{R}$. To see that $X'$ is a smooth submanifold $C^{k-1}$ we note that $X'$ is the set of zeros of $D_2f' : X\times TX\to\mathbb{R}$ in a neighborhood $N_\varepsilon(\Delta(X))\subset X\times TX$. Our assumption on the Hessian implies that these are regular values so $X'$ is $C^{k-1}$ submanifold since $D_2f'$ is $C^{k-1}$. That $X'$ is diffeomorphic to $X$ follows from the fact that $X'$ is $C^{k-1}$ close to $\Delta(X)$. This is immediate since $D_2f'$ is $C^{k-1}$ close to $D_2f$. □

For the remainder of this section $X$ will be a compact Riemannian manifold, $\Gamma$ will be a locally compact group with property $(T)$ and $K$ will be a fixed compact generating set, $\rho$ will be an isometric action of $\Gamma$ on $X$ and $\rho'$ will be a $C^k$ perturbation of $\rho$, where $k > 2$. We will denote by $\bar{\rho}'$ the $\Gamma$ action on $X\times X$ given by perturbing in the second factor: $\bar{\rho}'(\gamma)(x_1, x_2) = (\rho(\gamma)x_1, \rho'(\gamma)x_2)$. This induces actions on various spaces of functions which we also denote by $\bar{\rho}'$.

We now prove Theorem 1.1. We only give a proof with small loss of derivatives. The reader interested in lower regularity results depending only on Hilbert space techniques can produce a proof by combining this one with the proof in subsection 5.1.

*Proof of Theorem 1.1.* We first choose a function $f$ invariant under $\bar{\rho}$ as in Proposition 5.3. Given $\kappa > 0$, we choose $p$ with $\kappa < \frac{d}{p}$ where $d = \dim(X\times X) = 2\dim(X)$. We make a choice of $\eta > 0$, depending on $p$, to be specified below. We choose $U$ satisfying Lemma 4.5 for our choices of $p$ and $\eta$ and then Lemma 4.5 implies that there is a $\bar{\rho}'$ invariant function $f'$ with $\|f - f'\|_{p,k} < \eta$.

Proposition 4.3 implies that $\|f - f'\|_{k-\frac{d}{p}} < C_0\eta$ where $C_0$ depends only on $g$ and $p$. Choosing $\eta$ small enough and applying Proposition 5.4, we see that we have a submanifold $X'\subset X\times X$ which is diffeomorphic to $X$, $C^{k-\frac{d}{p}-1}$ close to $\Delta(X)$ and $\bar{\rho}'(\Gamma)$-invariant. The first two claims are contained in that proposition, the last follows from the definition $X' = \{(x, x')|f'(x')$ is the global minimum of $f'$ on $\{x\}\times X\}$. We let $p_i : X'\to X$ be the restriction to $X'$ of the projection $\pi_i : X\times X\to X$ on the $i$th factor where $i = 1, 2$. Note that each $\pi_i$ and therefore each $p_i$ is an equivariant map, where we view the first projection as to $X$ equipped with the action $\rho$ and the second as to $X$ equipped with the action $\rho'$. Since $X'$ is $C^{k-\frac{d}{p}-1}$ close to the diagonal, each $p_i$ is a $C^{k-\frac{d}{p}-1}$ diffeomorphism, and the map ${p_1}^{-1}\circ p_2$ is a $C^{k-\frac{d}{p}-1}$ small

diffeomorphism. Therefore ${p_1}^{-1}\circ p_2$ is a $C^{k-\frac{d}{p}-1}$ conjugacy between $\rho$ and $\rho'$. $\square$

## 6. Additional estimates and $C^{\infty,\infty}$ Local Rigidity

In this section, we prove a key lemmas on regularity in the context of isometric actions and their perturbations. From this, we deduce $C^{\infty,\infty}$ local rigidity in Theorem 1.1. In all that follows $\Gamma$ will be a locally compact group with property $(T)$ of Kazhdan and $K$ will be a fixed, compact generating set for $\Gamma$, containing a neighborhood of the identity in $\Gamma$. Furthermore, for simplicity of exposition, the letters $k$ and $l$ below always denote integers.

The strategy of the proof of the $C^{\infty,\infty}$ version of Theorem 1.1 is motivated by analogy with the iterative methods of $KAM$ theory but does not follow a $KAM$ algorithm, see Appendix D.1 for discussion.

**Remark:** In order to prove optimal regularity, we make use here of Corollary 2.8 and the resulting estimates in $L^p$ type Sobolev spaces. This allows us to give proofs that imply that, in the context of Theorem 1.1, a $C^\infty$ action $\rho'$ that is sufficiently $C^2$ close to an isometric action $\rho$ is conjugate back to $\rho$ by a $C^\infty$ map which is $C^{2-\kappa}$ small, for $\kappa$ depending on the $C^2$ size of the perturbation. The reader only interested in obtaining a $C^\infty$ conjugacy under some circumstances, rather than optimal circumstances, can easily modify the proofs to use only $L^2$ type Sobolev spaces and Theorem 2.1 instead.

We first state a proposition and lemma for isometric actions. We prove Proposition 6.1 from Lemma 6.2 and some results in subsection 5.1, and then use Proposition 6.1 to prove $C^{\infty,\infty}$ local rigidity for isometric actions. Lemma 6.2 will be proven later in this section. For notational convenience in the statement of this proposition and the proof of the $C^\infty$ case of Theorem 1.1, it is convenient to fix right invariant metrics $d_l$ on the connected components of $\mathrm{Diff}^l(X)$ with the additional property that if $\varphi$ is in the connected component of $\mathrm{Diff}^\infty(X)$, then $d_l(\varphi,\mathrm{Id})\leq d_{l+1}(\varphi,\mathrm{Id})$. To fix $d_l$, it suffices to define inner products $<,>_l$ on $\mathrm{Vect}^l(X)$ which satisfy $<V,V>_l\ \leq\ <V,V>_{l+1}$ for $V\in\mathrm{Vect}^\infty(X)$. Fixing a Riemannian metric $g$ on $X$, it is straightforward to introduce such metrics using the methods of section 4.

**Proposition 6.1.** *Let $X$ be a compact Riemannian manifold and $\rho$ be an isometric action of $\Gamma$ on $X$. Then for every integer $k\geq 2$ and every integer $l\geq k$ and every $\varsigma>0$ there is a neighborhood $U$ of the identity in $\mathrm{Diff}^k(X)$ such that if $\rho'$ is a $C^\infty$ action of $\Gamma$ on $X$ with $\rho(\gamma)^{-1}\rho'(\gamma)\in U$ for all $\gamma\in K$ then there exist a sequence $\psi_n\in\mathrm{Diff}^\infty(X)$ such that $\psi_n$*

*converge to a diffeomorphism* $\psi$ *in* $\mathrm{Diff}^l(X)$ *and* $\psi_n \circ \rho' \circ {\psi_n}^{-1}$ *converges to* $\rho$ *in* $C^l$ *and* $d_{k-1}(\psi_n, \mathrm{Id}) < \varsigma$ *for all* $n$.

This proposition is a consequence of the following lemma concerning regularity of invariant sections. We will only use this lemma for the trivial bundle $\xi = X \times \mathbb{R}^n$ equipped with the $\Gamma$ actions $\rho(\gamma)(x, v) = (\rho(\gamma)x, \sigma(\gamma)v)$ and $\rho'(\gamma)(x, v) = (\rho'(\gamma)x, \sigma(\gamma)v)$ where $\sigma : \Gamma \to O(n)$ is fixed, so we do not consider more general tensor bundles $\xi$. A similar statement is true in the general context, though one needs to replace $\mathrm{Diff}^k(X)$ in the statement with $\mathrm{Diff}^{k+1}(X)$.

**Lemma 6.2.** *Let* $\Gamma, X, \rho$ *be as in Proposition 6.1, let* $\xi = X \times \mathbb{R}^n$ *and let* $s$ *be a* $\rho(\Gamma)$ *invariant section of* $\xi$. *Then for every integer* $k \geq 2$, *every integer* $l \geq k$ *and every* $\eta > 0$ *there is a neighborhood* $U$ *of the identity in* $\mathrm{Diff}^k(X)$ *such that if* $\rho'$ *is a* $C^\infty$ *action of* $\Gamma$ *on* $X$ *with* $\rho(\gamma)^{-1}\rho'(\gamma) \in U$ *for all* $\gamma \in K$ *then the sequence* $s_n = \rho'(h)^n s$ *satisfies:*

(1) $\|s - s_n\|_{k-1} < \eta$ *for all* $n$ *and,*
(2) $s_n$ *converges in* $\mathrm{Sect}^l(X)$ *to a* $\rho'$ *invariant section* $s'$.

**Remarks:** We defer the proof of Lemma 6.2 until later in this section. Given a positive integer $l > k$, the proof of the lemma only requires that $\rho'$ is $C^{2l-k+1}$ rather than $C^\infty$. By shrinking $U$, it is possible to show the same result when $\rho'$ is $C^{l+1}$.

*Proof of Proposition 6.1.* The proof is very similar to the argument in subsection 5.1. We apply Proposition 5.1 to the action $\rho$, which produces a representation $\sigma : \Gamma \to \mathbb{R}^n$ and an equivariant embedding $s : X \to \mathbb{R}^n$. We let $\xi = X \times \mathbb{R}^n$ and define a action of $\Gamma$ on $\xi$ as specified before Lemma 6.2. Then $s$ is $\rho(\gamma)$ invariant for every $\gamma \in \Gamma$. Given $\eta > 0$, Lemma 6.2 implies that there is a neighborhood $U$ of the identity in $\mathrm{Diff}^k(X)$ such that for any action $\rho'$ with $\rho(\gamma)^{-1}\rho'(\gamma)$ in $U$ for all $\gamma \in K$ and the action $\rho'$ on $\xi$ defined before the statement of Lemma 6.2, we have that $s_n = \rho'(h)^n s$ satisfy $\|s - s_n\|_{k-1} < \eta$ and $s_n$ converge in $\mathrm{Sect}^l(X)$ to a $\rho'$ invariant section $s'$. It is clear that each $s_n$ is $C^\infty$. Choosing $\eta$ small enough and applying Lemma 5.2, we see that the maps $\psi_n = s^{-1} \circ \phi \circ s_n$ are $C^{k-1}$ small, $C^\infty$ diffeomorphisms of $X$, where $\phi$ is the normal projection from a neighborhood of $s(X)$ in $\mathbb{R}^n$ to $s(X)$. Letting $\psi = s^{-1} \circ \phi \circ s'$, it is clear that $\psi_n$ converge to $\psi$ in $\mathrm{Diff}^l(X)$ since $s_n$ converge to $s$ in $\mathrm{Sect}^l(\xi)$. That $\psi$ is a conjugacy between the actions $\rho'$ and $\rho$ follows as in the proof of Theorem 1.1 in subsection 5.1. $\square$

*Proof of* $C^{\infty,\infty}$ *local rigidity in Theorem 1.1.* If $\rho'$ is a $C^\infty$ perturbation of $\rho$, then there exists some $k > 1$, such that $\rho'$ is $C^k$ close to $\rho$.

We fix a sequence of positive integers $k = l_0 < l_1 < l_2 < \dots < l_i < \dots$ and will construct a sequence of $C^\infty$ diffeomorphisms $\phi_i$ such that the sequence $\{\phi_n \circ \dots \circ \phi_1\}_{n \in \mathbb{N}}$ converges in the $C^\infty$ topology to a conjugacy between $\rho$ and $\rho'$.

We let $\phi^i = \phi_i \circ \dots \circ \phi_1$ and $\rho_i = \phi^i \circ \rho' \circ (\phi^i)^{-1}$ and construct $\phi_i$ inductively such that

(1) $\rho_i$ is sufficiently $C^{l_i}$ close to $\rho$ to apply Proposition 6.1 to $\rho_i$ and $\rho$ with $l = l_{i+1}$ and $\varsigma = \frac{1}{2^{i+1}}$,
(2) $d_{l_i}(\phi_i, \mathrm{Id}) < \frac{1}{2^i}$ and,
(3) $d_{l_{i-1}}(\rho_i(\gamma) \circ \rho(\gamma)^{-1}, \mathrm{Id}) < \frac{1}{2^i}$ for every $\gamma \in K$.

Given $\phi^i$ and therefore $\rho_i$, we construct $\phi_{i+1}$. We have assumed that $\rho_i$ is close enough to $\rho$ in the $C^{l_i}$ topology to apply Proposition 6.1 with $l = l_{i+1}$ and $\varsigma = \frac{1}{2^{i+1}}$. Then we have a sequence of diffeomorphisms $\psi_n \in \mathrm{Diff}^\infty(X)$ such that $\psi_n \circ \rho_i \circ \psi_n{}^{-1}$ converges to $\rho$ in the $C^{l_{i+1}}$ topology and $d_{k-1}(\psi_n, \mathrm{Id}) < \frac{1}{2^{i+1}}$. We choose $n_i$ sufficiently large so that $\rho_{i+1} = \psi_{n_i} \circ \rho_{i-1} \circ \psi_{n_i}{}^{-1}$ is close enough to $\rho$ in the $C^{l_{i+1}}$ topology to apply Proposition 6.1 with $l = l_{i+2}$ and $\varsigma = \frac{1}{2^{i+2}}$ and so that $d_{l_i}(\rho_i(\gamma)\rho(\gamma)^{-1}, \mathrm{Id}) \leq \frac{1}{2^{i+1}}$ and then let $\phi_{i+1} = \psi_{n_i}$.

To start the induction it suffices that $\rho'$ is sufficiently $C^k$ close to $\rho$ to apply Proposition 6.1 with $l = l_1$ and $\varsigma = \frac{1}{2}$.

It remains to show that the sequence $\{\phi_n \circ \dots \circ \phi_1\}_{n \in \mathbb{N}}$ converges in the $C^\infty$ topology to a conjugacy between $\rho$ and $\rho'$. Combining condition (2) with the fact that $d_{l_i}(\phi_m, \mathrm{Id}) \leq d_j(\phi_m, \mathrm{Id})$ for all $j \geq l_i$, and the fact that $d_{l_i}$ is right invariant implies that $d_{l_{i-1}}(\phi_m, \mathrm{Id}) = d_{l_{i+1}}(\phi^m, \phi^{m-1}) \leq \frac{1}{2^m}$ for all $m \geq i$. This implies that $\{\phi^m\}$ is a Cauchy sequence in $\mathrm{Diff}^{l_i}(X)$ for all $i$, and therefore $\phi^m$ converge in $\mathrm{Diff}^\infty(X)$. Similarly, condition (3) implies $\rho_m$ converges to $\rho$ in the $C^\infty$ topology. □

**Remark:** The proof above can be made to work in a more general setting. Given an action $\rho$ such that for any large enough $k$ and any $l$ larger than $k$ and any action $\rho'$ which is sufficiently $C^k$ close to $\rho$, we can find a conjugacy between $\rho$ and $\rho'$ which is $C^l$ and $C^{k-n}$ small for a number $n$ which does not depend on $l$ or $k$, then we can use the method above to produce a $C^\infty$ conjugacy. More precisely, we need a bound on the $C^{k-n}$ size of the conjugacy that depends only on the $C^k$ size of the perturbation. To apply the argument in this setting, one produces a $C^l$ conjugacy $\varphi$ and then approximates it in the $C^l$ topology by a $C^\infty$ map $\tilde{\varphi}$ which will play the role of $\psi_n$ in the argument above. We use this argument to prove $C^{\infty,\infty}$ local rigidity in [FM2].

Before we proceed to prove Lemma 6.2, we need two additional estimates. Similar estimates are used in KAM theory. The first is a

convexity estimate on derivatives, which is also used in the proof of Hamilton's $C^\infty$ implicit function theorem, and which we take from [Ho]. To be able to prove a foliated variant of Lemma 6.2 below, we state these estimates in the context of foliated spaces. For the next two lemmas, let $(X, \mathfrak{F}, g_{\mathfrak{F}})$ be a foliated space equipped with a leafwise Riemannian metric as described in section 4. For our applications in subsection 7.3, it is important that $X$ need not be compact in either of the following lemmas.

**Lemma 6.3.** *Let $a, b, c$ be integers and $0 < \lambda < 1$ such that $c = a(1-\lambda) + b\lambda$ and let $f \in \mathrm{Sect}^k(\xi, \mathfrak{F})$. Then there is a constant $B$ depending only on $X, \mathfrak{F}$ and $g_{\mathfrak{F}}$ and $b$ such that:*

$$\|f\|_c \leq B\|f\|_a^{1-\lambda}\|f\|_b^\lambda$$

.

For $a, b, c$ not necessarily integral, this lemma is proven for functions on $\mathbb{R}^n$ in appendix A of [Ho]. This implies the proposition as stated by standard manipulations as in the proof of Proposition 4.3.

Given a collection elements $\phi_1 \ldots, \phi_n \in \mathrm{Diff}^\infty(X, \mathfrak{F})$ we require a certain type of bound on the norm of the composition $\phi_1 \circ \cdots \circ \phi_n$ as an operator on $k$-jets of tensors. Recall that we have a pointwise norm $\|j^k(\phi)(x)\|$ defined to be the operator norm of

$$j^k(\phi)(x) : J^k(X, \mathfrak{F})_x \to J^k(X, \mathfrak{F})_{\phi(x)}.$$

Then we can define the $k$ norm of $\phi$ by $\|\phi\|_k = \sup_X \|j^k(\phi)(x)\|$. Though we did not find the following precise estimate in the literature, this type of estimate is typical in KAM theory. In Appendix B, we give a proof that may be new, at least in that it makes no reference to coordinates.

**Lemma 6.4.** *Let $\phi_1, \ldots, \phi_n \in \mathrm{Diff}^k(X, \mathfrak{F})$. Let $N_k = \max_{1 \leq i \leq n} \|\phi_i\|_k$ and $N_1 = \max_{1 \leq i \leq n} \|\phi_i\|_1$. Then there exists a polynomial $Q$ depending only on the dimension of the leaves of the foliation and $k$ such that:*

$$\|\phi_1 \circ \ldots \circ \phi_n\|_k \leq N_1^{kn} Q(nN_k)$$

*for every $n \in \mathbb{N}$.*

This lemma has immediate consequences for the operator norms of $\rho'(h)$ on $C^k(X, \mathfrak{F})$ which we denote by $\|\rho'(h)\|_k$.

**Corollary 6.5.** *Under the assumptions of Lemma 6.2, for any $h \in \mathcal{U}(\Gamma)$, we have the following estimates:*

$$\|\rho'(h)^n\|_k \leq N_1^{kn} Q(nN_k)$$

*where $Q$ is the same polynomial as in Lemma 6.4 above and* $N_i = \max_{\mathrm{supp}(h)} \|\rho'(\gamma)\|_i$.

**Remark:** We require this estimate to be able to estimate the size of $\rho'(h)^n s$ in the $C^l$ topology, even when the group action $\rho'$ is only $C^k$ close to $\rho$ for some $k < l$. We do not know of another way to obtain such an estimate.

*Proof.* It follows from the definition that

$$\rho'(h)^n = \int_\Gamma h^{*n} \rho'(g) =$$

$$\int_\Gamma \cdots \int_\Gamma h(\gamma_1) \ldots h(\gamma_n) \rho'(\gamma_1) \ldots \rho'(\gamma_n).$$

One then applies Lemma 6.4 applied to each product of the form $\rho'(\gamma_1) \ldots \rho'(\gamma_n)$ and integrates. □

The polynomial $Q$ is computable in a straightforward manner for any given $k$ and dimension as follows easily from the proof.

*Proof of Lemma 6.2.* For the proof of this lemma, we let $p$ be such that $\frac{\dim(X)}{p} < 1$. By Lemma 4.5, for any $0 < C < 1$ and $F > 0$ and $\beta > 0$, we can choose a neighborhood $U$ of the identity in $\mathrm{Diff}^k(X)$ such that if $\rho'$ is a $C^\infty$ action with $\rho'(\gamma)\rho(\gamma)^{-1} \in U$ for all $\gamma \in K$, there exists $h \in \mathcal{U}(\Gamma)$ such that $\rho'(h)^n s$ converges to a $\rho'$ invariant section $s'$ where

$$\|\rho'(h)^n s - s'\|_{p,k} \leq \beta$$

for all $n$ and

$$\|\rho'(h)^{n+1} s - \rho'(h)^n s\|_{p,k} \leq C^n F.$$

Proposition 4.3 then implies that

$$\|\rho'(h)^{n+1} s - \rho'(h)^n s\|_{k-1} \leq C^n A F \tag{1}$$

and

$$\|\rho'(h)^n s - s'\|_{k-1} \leq A\beta$$

where $A$ depends only on $(X, g)$. The last inequality implies the first conclusion of Lemma 6.2 provided we chose $\beta < \frac{\eta}{A}$.

We will show that, possibly after shrinking $U$, $\rho'(h)^n s$ satisfies

$$\|(\rho'(h)^{n+1} s - \rho'(h)^n s)\|_l \leq C'^n P(nF_l) F \tag{2}$$

where $P$ is a fixed polynomial and $F_l = F_l(l) > 0$ and $0 < C' = C'(C, l) < 1$. This estimate immediately implies that $\rho'(h)^n s$ converges in $\mathrm{Sect}^l(X)$ so to prove the lemma it suffices to prove inequality (2).

We let $b = 2l - k + 1$ and define $F_l = \sup_{\mathrm{supp}(h)} \|\rho'(\gamma)\|_b$. We shrink $U$ so that $\|\rho'(\gamma)\|_1^b C < 1$ for every $\gamma \in \mathrm{supp}(h)$, let $C_h = \sup_{\mathrm{supp}(h)} \|\rho'(\gamma)\|_1$

and fix a constant $C'$ with $\sqrt{C_h^b C} < C' < 1$. Let $f_n = \rho'(h)^{n+1}s - \rho'(h)^n s$. Then Lemma 6.3 implies that

$$\|f_n\|_l \leq B\|f_n\|_{k-1}^{\frac{1}{2}}\|f_n\|_b^{\frac{1}{2}} \tag{3}$$

for $B$ depending only on $X$ and $b$. Inequality (1) provides a bound on $\|f_n\|_{k-1}$, so it remains to find a bound on $\|f_n\|_b$. Noting that $f_n = \rho'(h)^n(\rho'(h)s - s)$ Corollary 6.5 implies that

$$\|f_n\|_b \leq C_h^{nb} P(nF_l). \tag{4}$$

Inequality (2) is now immediate from inequalities (1), (3) and (4) and the definition of $C'$. □

**Remark on the choice of $U$:** There are two constraints on the choice of $U$:

(1) $U$ is small enough so that we can apply Lemma 4.5 as described in the first paragraph of the proof for $\beta < \frac{\eta}{A}$ and some $0 < C < 1$ and
(2) $U$ is small enough so that $\|\rho'(\gamma)\|_1^b C < 1$ for every $\gamma \in \mathrm{supp}(h)$.

It is easy to see that we can choose $U$ to satisfy these two conditions. An analogous remark applies to the proof of Lemma 7.7 below.

## 7. Foliated results

This section is devoted to the proof of Theorem 2.11. Though we can prove some special cases of Theorem 2.11 by the method of subsection 5.1, the general result requires that we use the method described in subsection 5.2 for isometric actions. We begin by recalling some facts about foliations and their holonomy groupoids.

7.1. **Holonomy groupoids and regular atlases.** We would like to be able to apply the definitions and results of section 4 to the "foliated space" defined by taking pairs of points on the same leaf of a foliation $\mathfrak{F}$ of $X$. There is a well-known difficulty in topologizing the set of pairs of points on the same leaf as a foliated space and it seems difficult even to make this space a measure space in a natural way without some additional assumption on the foliation. For product foliations $X = Y \times Z$ foliated by copies of $Z$, no difficulties occur and the space is simply $Y \times Z \times Z$. More generally, one usually considers the holonomy groupoid or graph of the foliation, which is a, possibly non-Hausdorff, foliated space. To avoid technical difficulties, we have assumed that our foliated spaces have Hausdorff holonomy groupoids.

We now briefly describe the holonomy groupoid $P$ of the foliated space $(X, \mathfrak{F})$ in order to define group actions on $P$ associated to $\rho$ and

$\rho'$. At each point $x$ in $X$, we fix a local transversal, $T_x$. Given a curve $c$ contained in a leaf $\mathfrak{L}_x$ of $\mathfrak{F}$, with endpoints $x$ and $y$, one can define the holonomy $h(c)$ of $c$ as the germ of the map from $T_x$ to $T_y$ given by moving along (parallel copies of) $c$. It is clear that $h(c)$ depends on the homotopy class of $c$. We can define an equivalence relation on paths $c$ from $x$ to $y$ by saying two paths $c$ and $c'$ are equivalent if $h(c) = h(c')$. Then $P$ is the set of equivalence classes of triples $(x, y, c)$ where $c$ is a curve joining $x$ to $y$ and two triples $(x, y, c)$ and $(x', y', c')$ are equivalent if $x = x', y = y'$ and $h(c) = h(c')$. There is an obvious topology on $P$ in which $P$ is a foliated space with leaves of the form $\mathfrak{L}_x \times \tilde{\mathfrak{L}}_x$ where $\mathfrak{L}_x$ is a leaf of $\mathfrak{F}$ and $\tilde{\mathfrak{L}}_x$ is the cover of $\mathfrak{L}_x$ corresponding to homotopy classes of loops at $x$ with trivial holonomy. When we wish to refer explicitly to the structure of $P$ as a foliated space, we will use the notation $(P, \tilde{\mathfrak{F}})$. As mentioned above, we will always assume that $P$ is Hausdorff in it's natural topology. There are two natural projections $\pi_1$ and $\pi_2$ from $P$ to $X$ defined by $\pi_1(x, y, c) = x$ and $\pi_2(x, y, c) = y$ both of which are continuous and leafwise smooth. A transverse invariant measure on $X$ defines one on $P$ and a leafwise volume form on $(X, \mathfrak{F})$ defines one on $(P, \tilde{\mathfrak{F}})$. Therefore, under the hypotheses of Theorem 2.11, we have a, possibly infinite, measure $\tilde{\mu}$ on $P$ defined by integrating the leafwise volume form against the transverse invariant measure. It is easy to see that $\tilde{\mu} = \int_X \tilde{\nu}_{\mathfrak{F}} d\mu$ where $\tilde{\nu}_{\mathfrak{F}}$ is the pullback of the leafwise volume form on leaves of $\mathfrak{F}$ to their holonomy coverings. For more detailed discussion, the reader should see either [CC] or [MS].

Given an action $\rho$ of $\Gamma$ on $(X, \mathfrak{F})$ defined by a homomorphism $\rho : \Gamma \to \mathrm{Diff}^k(X, \mathfrak{F})$ we can define an action $\rho_P$ on $P$ as follows. Take the diagonal action of $\rho$ on $X \times X$. This defines an action of $\Gamma$ on curves $c$ as above, which then descends to an action on $P$. It is immediate that if $\rho$ preserves $\mu$ then $\rho_P$ preserves $\tilde{\mu}$. If $\rho'$ is a $C^k$ foliated perturbation of $\rho$, then we can define an action $\rho'_P$ similarly, provided $\Gamma$ is compactly presented. We take the action on $X \times X$ defined by acting by $\rho$ on the first coordinate and $\rho'$ on the second. As long as $\rho(\gamma)$ is close to $\rho'(g)$, we can define $\rho'_P(\gamma)$ on $P$, since there is a canonical choice of a short, null homotopic, path from $\rho(\gamma)y$ to $\rho'(\gamma)y$ given by the length minimizing geodesic segment. Given a path $c$ from $x$ to $y$, we define $\rho'_P(\gamma)(x, y, c) = (\rho(\gamma)x, \rho'(g)y, c')$ where $c'$ is the concatenation of the path $\rho(g)c$ with the canonical path from $\rho(g)y$ to $\rho'(g)y$. Since $\Gamma$ is compactly presented, if $\rho$ is close enough to $\rho'$ it is easy to verify that lifting the generating set $K$ to $P$ defines an action $\rho'_P$ of $\Gamma$ on $P$.

It is immediate that $\pi_1 : (P, \rho'_P) \to (X, \rho)$ and $\pi_2 : (P, \rho'_P) \to (X, \rho')$ are equivariant. Note that compactness of $X$ implies that $\rho_P$ and $\rho'_P$ are close in the strong topology on $\mathrm{Diff}^k(P, \tilde{\mathfrak{F}})$.

When $P$ is Hausdorff, we can define a family of norms on sections of $J^k(P)$ by $\|f\|_p^p = \int_X \int_{\tilde{\mathfrak{L}}_x} \|f(x,y)\|^p d\nu_{\mathfrak{F}}(y) d\mu(x)$ where $\tilde{\mathfrak{L}}_x = {\pi_1}^{-1}(x)$. We can complete $\mathrm{Sect}(J^k(P))$ to a Banach space $L^p(J^k(P))$ of type $L^p_n$. Note that $C^k(P)$ is a linear subspace of $\mathrm{Sect}(J^k(P))$ and let $L^{p,k}(P, \mathfrak{F})$ be the closure of $F^k(P)$ in $L^p(J^k(P))$.

To obtain the required estimates for Theorem 2.11, we will also need estimates for the size of functions with respect to certain other $\Gamma$ invariant measures. Let $\lambda$ be any $\rho(\Gamma)$ invariant probability measure on $X$. Define an norm on $\mathrm{Sect}(J^k(P))$ by $\|f\|^p_{p,\lambda} = \int_X \int_{\mathfrak{L}_x} \|f\|^p d\nu_{\mathfrak{F}} d\lambda$. We can complete $\mathrm{Sect}(J^k(P))$ and $F^k(P)$ with respect to this norm to obtain Banach spaces $L^{p,\lambda}(J^k(P))$ and $L^{p,k,\lambda}(P, \mathfrak{F})$.

Except for the fact that we consider more general invariant measures and the corresponding function spaces, the following is a consequence of Propositions 4.1 and 4.2 above. The proofs of those propositions can be repeated almost verbatim to prove this one.

**Proposition 7.1.** *Let $\phi$ be a leafwise isometry of $(X, \mathfrak{F}, g_{\mathfrak{F}}, \mu)$.*

(1) *The maps $\phi_P$ on $L^p(J^k(P))$ and $L^{p,k}(P, \mathfrak{F})$ are isometric. Furthermore for any $\Gamma$ invariant probability measure $\lambda$, the maps $\phi_P$ on $L^{p,\lambda}(J^k(P))$ and $L^{p,k,\lambda}(P, \mathfrak{F})$ are isometric.*

(2) *For any $\varepsilon > 0$ and any $p_0 > 1$ there exists a neighborhood $U$ of the identity in $\mathrm{Diff}^k(X, \mathfrak{F})$ such that for any $(U, C^k)$-foliated perturbation $\phi'$ of $\phi$, the map $\phi'_P$ induces $\varepsilon$-almost isometries on $L^{p,k}(P, \mathfrak{F})$ and $L^p(J^k(P))$ and on $L^{p,\lambda}(J^k(P))$ and $L^{p,k,\lambda}(P, \mathfrak{F})$, for any $\Gamma$ invariant probability measure $\lambda$ on $X$ and any $p \leq p_0$.*

(3) *Let $f$ be a $\phi_P$ invariant compactly supported function in $C^k(P)$. Then for every $\delta > 0$ and every $p_0 > 1$ there exists a neighborhood $U$ of the identity in $\mathrm{Diff}^k(X, \mathfrak{F})$ such that any $(U, C^k)$-foliated perturbation $\phi'$ of $\phi$, we have that $\|\phi'_P f - f\|_{p,k} \leq \delta$ and $\|\phi'_P f - f\|_{p,k,\lambda} \leq \delta$ for every $\phi$ invariant probability measure $\lambda$ on $X$ and every $p \leq p_0$.*

(4) *If $H$ is a topological group and $\rho$ is a continuous leafwise isometric action of $H$ on $X$, then the actions of $H$ induced by $\rho_P$ on $L^{2,k}(P, \mathfrak{F})$ and on $L^{2,k,\lambda}(P, \mathfrak{F})$, for any $\Gamma$ invariant probability measure $\lambda$ on $X$, are continuous. Furthermore the same is true for any continuous action $\rho'$ which is an $(U, C^k)$-foliated perturbation of $\rho$.*

We need a proposition concerning covers of foliated spaces by certain kinds of foliated charts. This proposition follows from the proofs that any foliation can be defined by a *regular atlas* but since we require information not usually contained in the definition of a regular atlas, we sketch the proof here. For more discussion of regular atlases, see sections 1.2 and 11.2 of [CC]. We recall that for a foliated space $(X, \mathfrak{F})$ there is an associated metric space $Y$, such that there is a basis of foliation charts $(U, \phi)$ in $X$ of the form $\phi : U \to V \times B(0, r)$ where $U$ is an open in $X$, $V$ is an open set in $Y$ and $B(0, r)$ is a ball in $\mathbb{R}^n$.

**Proposition 7.2.** *Let $(X, \mathfrak{F}, g_{\mathfrak{F}})$ be a compact foliated space. Then there exists a positive number $r > 0$ and a finite covering of $X$ by foliated charts $(U_i, \phi_i)$ such that:*

1. *each $\phi_i : U_i \to V_i \times B(0, r)$ is a homeomorphism where $U_i \subset X$ and $V_i \subset Y$ are open and ${\phi_i}^{-1} : \{v_i\} \times B(0, r) \to \mathfrak{L} \cap U_i$ is isometric for all $v_i \in V_i$ and all $i$,*
2. *each $(U_i, \phi_i)$ is contained in a chart $(\tilde{U}_i, \tilde{\phi}_i)$ such that each $\tilde{\phi}_i : \tilde{U}_i \to V_i \times B(0, 2r)$ is a homeomorphism where $\tilde{U}_i \subset X$ and $V_i \subset Y$ are open and ${\phi_i}^{-1} : \{v_i\} \times B(0, 2r) \to \mathfrak{L} \cap \tilde{U}_i$ is isometric for all $v_i \in V_i$ and all $i$*

*Proof.* Let $\mathcal{W}$ be any maximal foliated atlas for $(X, \mathfrak{F})$. Since $X$ is compact, we can choose a finite cover of $X$ by $(W_j, \psi_j)_{1 \leq j \leq k} \subset \mathcal{W}$. Let $\eta$ be a Lebesgue number for the cover of $X$ by $W_j$, i.e. $B(x, \eta)$ is entirely contained in one $W_j$ for every $x \in X$. Let $2d$ be the largest number so that $B(x, \eta)$ contains a chart of the form $(W_i, \psi_i)$ in $\mathcal{W}$ such that $\psi_i(W_i) = V_i \times B(0, 2d)$ and $\psi_i|_{\{v_i\} \times B(0,2d)}$ is isometric for every $v_i \in V_i$. Let $U_i$ be ${\phi_i}^{-1}(V_i \times B(0, d)$ and let $\phi_i = \psi_i|_{U_i}$. Clearly the charts $(U_i, \phi_i)$ satisfy the conclusions of the proposition. Since $X$ is compact, we can pick a finite subset of these charts that cover $X$. □

7.2. **Proof of Theorem 2.11.** In this subsection we prove Theorem 2.11. The approach is based on the proof of Theorem 1.1 from subsection 5.2. In place of working on functions on $X \times X$ we work with functions in the spaces $C^k(P)$ and $L^{p,k}(P, \tilde{\mathfrak{F}})$ defined in subsection 7.1. By Theorem 2.4, we can assume without loss of generality that $\Gamma$ is compactly presented. Let $\rho$ be the leafwise isometric action specified in Theorem 2.11 and $\rho'$ the $C^k$ foliated perturbation of $\rho$. Then we have two $\Gamma$ actions $\rho_P$ and $\rho'_P$ on $P$ as defined in the last section. As before we will start with an invariant function $f$ for $\rho_P$, in this case a compactly supported function in $C^k(P)$ and therefore $L^{p,k}(P, \tilde{\mathfrak{F}})$, and construct the desired conjugacy from a $\rho'_P$ invariant function $f'$ close to $f$ in $L^{p,k}(P, \tilde{\mathfrak{F}})$. We construct $f$ in a manner analogous to Proposition

5.3. We first define a subset $\Delta \subset P$ which is the set of $\{(x, x, c_x) | x \in X\}$ where $c_x$ is the constant loop at $x$. Given a point $x \in X$, we denote by $\Delta(x) = (x, x, c_x)$. Note that this defines canonically a point $\tilde{x}$ in $\tilde{\mathfrak{L}}_x$, since the leaf of $\tilde{\mathfrak{F}}$ in $P$ through $\Delta(x)$ is $\mathfrak{L}_x \times \tilde{\mathfrak{L}}_x$ and $\tilde{x}$ is the projection of $\Delta(X) \cap (\mathfrak{L}_x \times \tilde{\mathfrak{L}}_x)$ to $\tilde{\mathfrak{L}}_x$. Given a point $x$ in $X$, we will refer to $B_{\mathfrak{F}}(x, r)$ as the ball of radius $r$ about $x$ in the leaf through $\mathfrak{L}_x$ and $B_{\tilde{\mathfrak{F}}}(\Delta(x), r)$ for the ball of radius $r$ about $\Delta(x)$ in $\mathfrak{L}_x \times \tilde{\mathfrak{L}}_x \subset P$. Recall that $N(x)$ is the radius of the largest normal ball containing $x$ in $\mathfrak{L}_x$, and that $N(x)$ is bounded below by a positive number since $X$ is compact. We sometimes write coordinates on $P$ as $p = (\pi_1(p), y)$ where $y \in {\pi_1}^{-1}(\pi_1(p))$.

**Proposition 7.3.** *Let $\rho$ be a leafwise isometric action of a group $\Gamma$ on a compact foliated space $(X, \mathfrak{F}, g_{\mathfrak{F}})$, then the action $\rho_P$ leaves invariant any function on $P$ which is a function of $d_{\mathfrak{L}_x}(\pi_1(p), \pi_2(p))$ or $d_{\tilde{\mathfrak{L}}_x}(\widetilde{\pi_1(p)}, y)$. Let $r > 0$ be as in Proposition 7.2 and such that $N(x) > 2r$ for all $x \in X$, then for any $0 < \varepsilon \leq r$, there exists a compactly supported invariant function $f \in C^\infty(P)$ such that:*

1. *$f$ takes the value $1$ on $\Delta$,*
2. *$f \geq 0$ and $f(p) = 0$ if $\pi(p) = x$ and $p \notin B_{\tilde{\mathfrak{F}}}(\Delta(x), 2\varepsilon)$,*
3. *$f(p) < 1/2$ if $\pi(p) = x$ and $p \notin B_{\tilde{\mathfrak{F}}}(\Delta(x), \varepsilon)$*
4. *the Hessian of $f$ restricted to $\pi^{-1}(x)$ is negative definite on the closure of $B_{\tilde{\mathfrak{F}}}(\Delta(x), \varepsilon) \cap \pi^{-1}(x)$*

*Proof.* The proof is very similar to the proof of Proposition 5.3.

To define $f$ we first define $f_{x_0} : \tilde{\mathfrak{L}}_x \to \mathbb{R}$ by $1 - \frac{1}{2\varepsilon^2} d(x_0, x)^2$. Our assumptions on $\varepsilon$ imply that $B(x_0, 2\varepsilon)$ is contained in a normal neighborhood of $x_0$ and so $d(x, x_0)^2$ is a smooth function of $x$ and $x_0$ inside $B(x_0, 2\varepsilon)$, see for example [KN, IV.3.6]. It is clear that $f(x_1, x_2) = f_{\tilde{x}_1}(x_2)$, satisfies $1, 3$ and 4, but it may not be smooth, fails to satisfy 2 and may not be compactly supported. Modifying $f_x$ outside $B(x, \varepsilon)$ produces a smooth,positive, compactly supported function satisfying all the above conditions. This is easily done by choosing any smooth function $g : \mathbb{R} \to \mathbb{R}$ such that $g$ agrees with $1 - \frac{1}{2\varepsilon^2} x^2$ to all orders for all $x \leq \varepsilon$ and $g < 1/2$ for all $x \geq \varepsilon$ and $g = 0$ for all $x \geq 2\varepsilon$. We then let $f(x_1, x_2) = g(d_{\tilde{\mathfrak{L}}_x}(\tilde{x}_1, x_2))$ where $x_1 = \pi_1(p)$ and $x_2$ is the coordinate of $p$ in $\tilde{\mathfrak{L}}_x = {\pi_1}^{-1}(x_1)$. □

**Remark:** We need $f$ to be compactly supported on ${\pi_1}^{-1}(x)$ for all $x \in X$. For this reason we choose $f$ with a global maximum along $\Delta(X)$ rather than a minimum as in Proposition 5.3.

One can now produce an invariant function $f'$ in $L^{2,k}(P,\tilde{\mathfrak{F}})$ exactly as in the proof of Theorem 1.1 in subsection 5.2. In order to control the behavior of $f'$ in $L^{2,k,\lambda}(P,\tilde{\mathfrak{F}})$ for any $\Gamma$ invariant probability measure $\lambda$ on $X$ as well, we need to produce $f'$ using Theorem 2.1 rather than Theorem 1.6. The following is a sharpening of Lemma 4.5, and as in the proof of that lemma, we use Theorem 2.1 in conjunction with Corollary 2.8 and Lemma 4.4 to obtain optimal regularity.

**Lemma 7.4.** *Let $\rho$ and $\Gamma$ be as in Theorem 2.11 and $f$ be a compactly supported $\rho_P(\Gamma)$ invariant function in $C^\infty(P)$. Given constants $\varsigma_1 > 0, F > 0, 0 < C < 1$ and $p{\geq}2$ there exists a neighborhood $U$ of the identity in $\mathrm{Diff}^k(X,\mathfrak{F})$ and a function $h{\in}\mathcal{U}(\Gamma)$ such that if $\rho'$ is any $(U,C^k)$-foliated perturbation of $\rho$:*

1. *$\rho'_P(h)^n f$ converges pointwise almost everywhere to $\rho'_P$ invariant function $f'$,*
2. *$\|f-f'\|_{p,k} < \varsigma_1$ and $\|f-f'\|_{p,k,\lambda} < \varsigma_1$ for every $\rho(\Gamma)$ invariant probability measure $\lambda$ on $X$,*
3. *$\|\rho'_P(h)^{n+1}f - \rho'_P(h)^n f\|_{p,k,\lambda} < C^n F$ for every $\rho(\Gamma)$ invariant probability measure $\lambda$ on $X$.*

*Proof.* Given $\varsigma_1 > 0$ and $p{\geq}2$, we choose $\varepsilon > 0$ satisfying the hypotheses of Theorem 2.1 and 2.8 and $\delta > 0$ depending on $\varsigma_1$ to be specified below. Choosing $U$ small enough, Proposition 7.1 implies that $\mathrm{disp}_K(f) < \delta$ for the $\rho'_P$ action on $L^{p,k,\lambda}(P,\mathfrak{F})$ and $L^{2,k,\lambda}(P,\mathfrak{F})$ for every $\rho(\Gamma)$ invariant probability measure $\lambda$ and also that $\rho'_P(k)$ is an $\varepsilon$-almost isometry on $L^{2,k,\lambda}(J^k(P))$ and $L^{p,\lambda}(J^k(P))$ for every $k{\in}K$ and every $\rho_P$ invariant probability measure $\lambda$ and also that $\rho'_P$ is a continuous action on all of these spaces. Fix a constant $0 < C < 1$ and a function $h{\in}\mathcal{U}(\Gamma)$ so that Theorem 2.1 and Corollary 2.8 are both satisfied for $h$ and $C$. Each theorem yields a constant $M_2$ and $M_p$ and we let $M = \max(M_2, M_p)$. Note that $M$ and $h$ depend only on $\Gamma, K, p$ and $C$ and some function $f{\in}\mathcal{U}(\Gamma)$. Therefore we can choose $\delta$ such that $\frac{MC}{1-C}\delta < \varsigma_1$ and $\delta leq \frac{F}{M}$. Then the sequence $\{\rho'_P(h^n)f\}$ satisfies $\|\rho'_P(h^n)f - \rho'_P(h^{n-1})f\|_{p,\lambda}{\leq}MC^n\delta{\leq}C^nF$ and $\|\rho'_P(h^n)f - \rho'_P(h^{n-1})f\|_{2,k,\lambda}{\leq}MC^n\delta{\leq}C^nF$ and $K$-displacement of $\rho'_P(h^n)f$ is less than $C^n\delta$ in $L^{2,k,\lambda}(P,\mathfrak{F})$ for every $\rho(\Gamma)$ invariant measure $\lambda$ on $X$. By Lemma 4.4, this implies that $\rho'_P(h)^n f$ converges pointwise $\lambda$ almost everywhere to a $\rho'_P$ invariant function $f'$ which is in $L^{p,k,\lambda}(P,\mathfrak{F})$. Furthermore, our choice of $\delta$ implies that $\|f-f'\|_{p,k,\lambda} < \varsigma_1$ for every $\Gamma$ invariant measure $\lambda$ on $X$. □

To obtain control over $f'$ and the resulting conjugacy on a set $S$ as described in Theorem 2.11, we need to consider a certain class of $\Gamma$ invariant measures on $X$. Let $\mu = \int_{\mathcal{P}(X)} \mu_e d\bar{\mu}(e)$ be an ergodic decomposition

for $\mu$, where each $\mu_e$ is a $\Gamma$ ergodic measure on $X$ and $\bar{\mu}$ is a measure on the space $\mathcal{P}(X)$ of probability measures on $X$ supported on the $\Gamma$ ergodic measures. Let $\mathcal{P}(X)$ be the space of regular Borel probability measures on $X$ and define a Markov operator $M : X \to \mathcal{P}(X)$ be letting $M_x = \frac{1}{\nu_{\mathfrak{F}}(B_{\mathfrak{F}}(x,r))}\nu_{\mathfrak{F}}|_{B_{\mathfrak{F}}(x,r)}$. Then $M$ defines an operator on continuous functions on $X$ by $Mg(x) = \int_X g(y)dM_x(y)$ for $g$ in $C^0(X)$ and dually an operator on $\mathcal{P}(X)$ by $M\nu(f) = \int_X Mf d\nu = \int_X \int_X f(y)dM_x(y)d\nu(x)$ for $\nu \in \mathcal{P}(X)$. Note that $M$ commutes with elements of $\mathrm{Diff}^k(X, \mathfrak{F})$ which are leafwise isometric. This implies that for any $\rho(\Gamma)$ invariant probability measure $\nu$ on $X$, the probability measure $M\nu$ is also $\Gamma$ invariant. We will be particularly interested in measures of the form $M\mu_e$ where $\mu_e$ is an ergodic component of $\mu$.

**Lemma 7.5.** *Let $f \in C^k(P)$ be compactly supported and let $f'$ be a function which is in $L^{p,k,\lambda}(P, \tilde{\mathfrak{F}})$ for every $\rho(\Gamma)$ invariant probability measure $\lambda$ on $X$, and such that $\|f - f'\|_{p,k,\lambda} \leq A$ for every $\lambda$. Then for any $\lambda$ and $\lambda$ almost every $x$, the restriction of $f'$ to ${\pi_1}^{-1}(x)$ is in $C^{k-\frac{d}{p}}$. Furthermore, there exist constants $C$ and $r$ depending only on $(X, \mathfrak{F})$ and a set $S \subset X$ depending on $f'$ such that $\lambda(S) > 1 - \sqrt{A}$ and*

$$\|(f - f')|_{{\pi_1}^{-1}(B_{\mathfrak{F}}(x,r))}\|_{k-\frac{d}{p}} \leq C\sqrt{A}$$

*for every $x \in S$.*

*Proof.* To see the first claim, we consider the measure $M\lambda$. The fact that $\|f - f'\|_{p,k,M\lambda} < A$ implies that $\|f'\|_{p,k,M\lambda}$ is finite. Applying the definition of $M$, this means that $\int_{B_{\mathfrak{F}}(x,r)} \int_{{\pi_1}^{-1}(x)} \|j^k(f')\|^p d\nu_{\mathfrak{F}} d\nu_{\mathfrak{F}}$ is finite for $\lambda$ almost every $x$. Then Proposition 4.3 implies that $f'$ is $C^{k-\frac{d}{p}}$ on ${\pi_1}^{-1}(B_{\mathfrak{F}}(x,r))$.

Let $v_0 = \min_{x \in X}(\nu_{\mathfrak{F}}(B_{\mathfrak{F}}(x,r)))$ and define $S$ to be the set of $x$ where

$$\int_{B_{\mathfrak{F}}(x,d)} \int_{\pi^{-1}(x)} \|j^k(f) - j^k(f')\|_k^p d\nu_{\mathfrak{F}} d\nu_{\mathfrak{F}} \leq v_0\sqrt{A}. \tag{5}$$

We first verify that $\lambda(S) \geq 1 - \sqrt{A}$ for every $\lambda$. We are assuming that $\|f - f'\|_{2,k,\lambda} \leq A$ for every $\rho$ invariant probability measure $\lambda$ on $X$. By definition of $L^{2,k,M\lambda}(P, \tilde{\mathfrak{F}})$ this means that

$$\int_X \int_{{\pi_1}^{-1}(x)} \|j^k(f) - j^k(f')\|^p d\nu_{\mathfrak{F}} dM(\lambda) =$$

$$\int_X \int_X \int_{{\pi_1}^{-1}(x)} \|j^k(f) - j^k(f')\|^p d\nu_{\mathfrak{F}} dM(x) d\lambda \leq A.$$

This implies that

$$\frac{1}{\nu_{\mathfrak{F}}(B_{\mathfrak{F}}(x,r))}\int_{B_{\mathfrak{F}}(x,r)}\int_{\pi_1{}^{-1}(x)}\|j^k(f)-j^k(f')\|^p d\nu_{\mathfrak{F}}d\nu_{\mathfrak{F}}<\sqrt{A}$$

or

$$\int_{B_{\mathfrak{F}}(x,r)}\int_{\pi_1{}^{-1}(x)}\|j^k(f)-j^k(f')\|_k^p d\nu_{\mathfrak{F}}d\nu_{\mathfrak{F}}\leq\sqrt{A}\nu_{\mathfrak{F}}(B_{\mathfrak{F}}(x,r))$$

on a set of $\lambda$ measure at least $1-\sqrt{A}$. Therefore the set $S$ defined by equation 5 has $\lambda$ measure at least $1-\sqrt{A}$ as desired.

Proposition 4.3 then implies that

$$\|(f-f')|_{\pi_1{}^{-1}(B_{\mathfrak{F}}(x,r))}\|_{k-\frac{d}{p}}<C'v_0\sqrt{A}$$

for every $x\in S$, where $C$ is a constant depending only on $X,\mathfrak{F}$ and $g_{\mathfrak{F}}$ and letting $C=C'v_0$ completes the proof. □

Fix $p$ with $\frac{\dim(\mathfrak{L}_x)}{p}<\kappa$ and fix a function $f$ as in the conclusion of Proposition 7.3 with $\varepsilon=r/2$ for the remainder of this section. We choose a constant $\varsigma_1$ to be specified below and let $f'$ be the function produced by Lemma 7.4. Then Lemma 7.5 combined with the definition of $f$ and $S$ implies that for every $x\in S$:

(1) $f'(p)<C\sqrt{\varsigma_1}$ if $\pi_1(p)=x$ and $p\notin B_{\tilde{\mathfrak{F}}}(\Delta(x),r)$,
(2) $f'(p)<1/2+C\sqrt{\varsigma_1}$ if $\pi_1(p)=x$ and $p\notin B_{\tilde{\mathfrak{F}}}(\Delta(x),\frac{r}{2})$
(3) for every $y\in B_{\mathfrak{F}}(x,r)$, the Hessian of $f'$ restricted to $\pi_1{}^{-1}(y)$ is negative definite on $B(\Delta(y),\frac{r}{2})\cap\pi_1{}^{-1}(y)$

Choosing $\varsigma_1\leq\frac{1}{100C^2}$ so $C\sqrt{\varsigma_1}\leq\frac{1}{10}$ this implies that for $x\in X$, the function $f'$ has a maximum on $\pi_1{}^{-1}(x)$ at a point $\tilde{\phi}(x)$ where the value is at least $\frac{9}{10}$ and that this maximum is the only local maximum with value greater than $\frac{6}{10}$. Since $f'$ is invariant under $\rho'_P$ it follows that if $\rho(\gamma)(x)\in S$ then $\tilde{\phi}(\rho(g)(x))=\rho'(g)\tilde{\phi}(x)$ since both points will be the global maxima of $f'$ on $\pi_1{}^{-1}(\rho(g)x)$. Furthermore, it follows by $\rho'_P(\Gamma)$ invariance of $f'$ that for every $x\in\Gamma{\cdot}S$, there is a unique global maximum for $f'|_{\pi_1{}^{-1}(x)}$. Therefore we can define the conjugacy between $\rho$ and $\rho'$ on a set of full measure in $X$ by letting $\tilde{\phi}(x)$ be the unique global maximum of $f'$ on the fiber $\pi_1{}^{-1}(x)=\tilde{\mathfrak{L}}_x$ and letting $\phi(x)=\pi_2(\tilde{\phi}(x))$.

We remark that it is possible to show that $d_{\mathfrak{L}_x}(x,\phi(x))^2\leq C'\sqrt{\varsigma_1}r^2$ for all $x\in S$ directly from the definition of $f,f'$ and $\phi$, where $C'=\frac{C}{v_0}$ is as in the proof of Lemma 7.5.

In order to make the following more readable, we let $k'=k-\frac{d}{p}-1$. We now show the map $\phi$ is leafwise $C^{k'}$ for $x\in S$, and therefore, by equivariance, leafwise $C^{k'}$ almost everywhere. Consider $x\in S$. We will show that $\phi$ is $C^{k'}$ and $C^{k'}$ close to the identity on $B_{\mathfrak{F}}(x,r)$.

Note that ${\pi_1}^{-1}(B_{\mathfrak{F}}(x,r)))$ is diffeomorphic to $B_{\mathfrak{F}}(x,r)\times\tilde{\mathfrak{L}}_x$. Let $D_2f' : B_{\mathfrak{F}}(x,r)\times T(\tilde{\mathfrak{L}}_x)\to\mathbb{R}$ be the derivative of $f'$ in the second variable. Let $N(\frac{r}{2}, B_{\mathfrak{F}}(x,r))$ be the $\frac{r}{2}$ neighborhood of $\Delta(B_{\mathfrak{F}}(x,r))$ in $B_{\mathfrak{F}}(x,r)\times\tilde{\mathfrak{L}}_x$ and $TN(\frac{r}{2}, B_{\mathfrak{F}}(x,r))$ the restriction of the bundle $\mathfrak{L}_x\times T\tilde{\mathfrak{L}}_x$ to that set. If $x\in S$ then the set of points $(x,\tilde{\phi}(x))$ is $D_2f'^{-1}(0)\cap TN(\frac{r}{2}, B_{\mathfrak{F}}(x,r))$ and 0 is a regular value of $D_2f'$, since the Hessian is negative definite on $\tilde{\mathfrak{L}}_{\tilde{y}}\cap B(\tilde{y},\frac{r}{2})$ for every $y\in B_{\mathfrak{F}}(x,r)$. This implies that the set $(x,\tilde{\phi}(x))\subset N(\frac{r}{2}, B_{\mathfrak{F}}(x,r))\subset TN(\frac{r}{2}, B_{\mathfrak{F}}(x,r))$ is a $C^{k'}$ submanifold and so $\tilde{\phi}$ is $C^{k'}$ on $B_{\mathfrak{F}}(x,r)$. This implies that $\phi$ is $C^{k'}$ on $B_{\mathfrak{F}}(x,r)$. Since $f'$ is $C^{k'}$ close to $f$ on ${\pi_1}^{-1}(B_{\mathfrak{F}}(x,r))$, the functions $D_2f$ and $D_2f'$ are $C^{k'}$ close on $TN(\frac{r}{2}, B_{\mathfrak{F}}(x,r))$. This implies that the submanifolds $D_2f^{-1}(0)$ and $D_2f'^{-1}(0)$ are $C^{k'}$ close, which then implies that $\phi$ is $C^{k'}$ close to the identity on $B_{\mathfrak{F}}(x,r)$. More precisely by choosing $\varsigma_1$ small enough, we can assume that the $C^{k'}$ norm of $\phi-\mathrm{Id} : B_{\mathfrak{F}}(x,r)\to B_{\mathfrak{F}}(x,2r)$ is less than $\varsigma$. We let $\bar{\varsigma}_1$ be the value of $\varsigma_1$ required for this estimate and let $\varsigma_1 = \min(\bar{\varsigma}_1, \frac{1}{100C^2v_0^2}, \sqrt{\varsigma})$. We let $U_0\subset \mathrm{Diff}^k(X,\mathfrak{F})$ be the neighborhood of the identity such that for any $(U_0,C^k)$-foliated perturbation $\rho'$ of $\rho$, the function $f'$ produced by Proposition 5.4 satisfies $\|f-f'\|_{p,k,\lambda}\leq\varsigma_1$ for every $\rho'(\Gamma)$ invariant measure $\lambda$ on $X$. Then we have verified conclusions (1), (2) and (3) of Theorem 2.11 for any $(U_0,C^k)$-foliated perturbation $\rho'$ of $\rho$.

To show the final estimate in the statement of Theorem 2.11, we need the following (well-known) quantitative refinement of [M, III.5.12].

**Lemma 7.6.** *There exists a constant $0<t<1$, depending only on $\Gamma$ and $K$, such that for any $0<\eta<\frac{1}{2}$ and any ergodic action of $\Gamma$ on a finite measure space $(X,\mu)$ and any set $S$ of measure $1-\eta$, there is $k$ in $K$ such that $\mu((kS\cup S)^c)\leq t\eta$.*

*Proof.* Assume not. Then for all $t$ with $0<t<1$, there exists $S$ of measure $1-\eta$ such that $\mu(kS\cup S)\leq 1-t\eta$ for all $k\in K$. We will use this fact to show that the characteristic function $\chi_S$ has $K$-displacement $(1-t)\eta$ and use this to produce a $\Gamma$ invariant function which is closer to the characteristic function of $S$ than any constant function.

Since $\mu(kS)+\mu(S)-\mu(kS\cap S)=\mu(kS\cup S)$ and $\mu(S)=\mu(kS)=1-\eta$, we have $\mu(kS\cap S)\geq(1-\eta)-(1-t)\eta$. Therefore $\mathrm{disp}_K(\chi_S)\leq\sqrt{(1-t)2\eta}$ in $L^2(X,\mu)$. By the standard linear analogue of Theorem 1.6 (which is an easy consequence of Lemma 3.4) there is a constant $C$ depending only on $\Gamma$ and $K$ and a $\Gamma$ invariant function within $C\sqrt{(1-t)2\eta}$ of $\chi_S$.

Since the orthogonal complement of the constant functions are the functions of integral zero, the distance from $\chi_S$ to the constant functions is $\sqrt{\eta(1-\eta)}$. Since $1-\eta > \frac{1}{2}$, we have a contradiction provided $C\sqrt{(1-t)2\eta} < \frac{\sqrt{\eta}}{\sqrt{2}}$ or $2C\sqrt{1-t} < 1$. So for $t > 1 - \frac{1}{4C^2}$ we are done. $\square$

Since $\mu_e(S) \geq 1 - \varsigma$ for some $\varsigma > 0$ for almost every ergodic component $\mu_e$ of $\mu$, it follows from Lemma 7.6 that the measure of $S_n = K^n{\cdot}S \cup \cdots K{\cdot}S \cup S$ is at least $1 - t^n\varsigma$.

Choose a neighborhood of the identity $U_1 \subset \mathrm{Diff}^k(X,\mathfrak{F})$ such that for every $x \in X$ and every $\gamma \in K$ we have $\|j^k(\rho'(\gamma))(x)\| \leq (1+\varsigma)$ for any $(U_1, C^k)$-foliated perturbation $\rho'$ of $\rho$. Let $U = U_1 \cap U_0$ and let $\rho'$ be $(U, C^k)$-foliated perturbation of $\rho$. Then the fact that $\|j^{k'}(\phi)(x)\| < 1+\varsigma$ for every $x \in S$, combined with the chain rule, the definition of $U_0$ and the fact that $\mu(S_n^c) < t^n\varsigma$, imply conclusion (4) of the theorem.

The remaining claim of the theorem states that given a positive integer $l$ then, if $U$ is small enough, $\phi$ is $C^l$. To prove this claim, it clearly suffices to see that $f'$ is $C^{l+1}$ on ${\pi_1}^{-1}(B_{\mathfrak{F}}(x,r))$ for almost every $x$. This is exactly the content of Lemma 7.7 in the next subsection.

7.3. **Improving regularity of $\varphi$ in Theorem 2.11.** We retain all notations and conventions from the previous two subsections. As remarked at the end of the last subsection, to complete the proof of Theorem 2.11, it suffices to prove that $f'$ is $C^{l+1}$ on ${\pi_1}^{-1}(B_{\mathfrak{F}}(x,r))$ for almost every $x$ in $X$. This is exactly the content of the following lemma.

**Lemma 7.7.** *Let $f$ in $C^k(P)$ be a compactly supported, $\rho_P(\Gamma)$ invariant function. For any $k \geq 3$, given a positive integer $l \geq k$, there exists $U \in \mathrm{Diff}^k(X,\mathfrak{F})$ such that for $\rho'$ a $(U, C^k)$-foliated perturbation of $\rho$ defined by a map from $\Gamma$ to $\mathrm{Diff}^{2l-k+1}(X,\mathfrak{F})$, the sequence $\{\rho'_P(h)^n f\}$ converges pointwise almost everywhere to a (measurable) function $f'$ on $P$ such that for almost every $x \in X$, the restriction of $f'$ to ${\pi_1}^{-1}(B_{\mathfrak{F}}(x,r))$ is $C^l$.*

To simplify the argument, we will use the operator $M$ on $\mathcal{P}(X)$ defined in subsection 7.2 and consider the measure $M\mu$.

We also introduce another technical mechanism to simplify the proof. Given $h \in \mathcal{U}(\Gamma)$, we can define a measure on $\Gamma$ by $h\mu_\Gamma$ where $\mu_\Gamma$ is Haar measure on $\Gamma$. We can then define a probability space $\Omega = \prod_{\mathbb{Z}} \Gamma$ with measure $\lambda = \prod_{\mathbb{Z}} h\mu_\Gamma$ and the left shift is an invertible measure preserving transformation $T$ of $(\Omega, \lambda)$. For any measure preserving action $\sigma$ of $\Gamma$ on a space $Y$, we can define a skew product extension

by $T_\sigma(\omega, y) = (T(\omega), \sigma(\omega_0)y)$ and ${T_\sigma}^{-1}(\omega, y) = (T^{-1}(\omega), \sigma(\omega_{-1})^{-1}y)$. Identifying functions on $Y$, or more generally, sections of bundles over $Y$, with their pullbacks to $\Omega \times Y$, it is clear that $\int_\Omega T_\sigma f d\lambda = \rho(h)f$ for every function $f$ on $Y$.

Before proving Lemma 7.7, we state the variant of Corollary 6.5 needed here. Since $(P, \mathfrak{F})$ is a foliated space, we can use the definitions of norms on $\mathrm{Diff}^k(X, \mathfrak{F})$ from section 6 to define norms on $\mathrm{Diff}^k(P, \mathfrak{F})$ and the estimates from Lemma 6.4 clearly hold for maps of $P$ as well. If $\psi(x, y) = (\phi_1(x), \phi_2(y))$, then it follows from the definitions that $\|\psi\|_k = \max(\|\phi_1\|_k, \|\phi_2\|_k)$. If $\phi_1$ is a $C^k$ leafwise isometry and $\phi_2$ is $(U, C^k)$-foliated perturbation of $\phi_1$, and we let $\psi = (\phi_1, \phi_2)$, then $\|\psi\|_k = \|\phi_2\|_k$. Similarly for $h \in \mathcal{U}(\Gamma)$ we can define the operator norm of $\rho'_P(h)$ acting on $J^k(P, \mathfrak{F})$ which we denote by $\|\rho'_P(h)\|_k$.

**Corollary 7.8.** *Under the assumptions of Lemma 7.7, for any function $h \in \mathcal{U}(\Gamma)$ we have the following estimate:*

$$\|\rho'_P(h)^n\|_k \leq N_1^{kn} Q(nN_k)$$

*where $Q$ is the same polynomial as in Lemma 6.4 above and $N_i = \max_{\mathrm{supp}(h)} \|\rho'(\gamma)\|_i$.*

**Remarks:**

(1) The proof is identical to the proof of Corollary 6.5, so we omit it.
(2) The fact that we need only consider $\|\rho'(\gamma)\|_i$ and not $\|\rho'_P(\gamma)\|_i$ in the statement of the corollary follows from the fact that $\rho'$ is a $(U, C^k)$ foliated perturbation of $\rho$ which implies $\|\rho' g)\|_i = \|\rho'_P(\gamma)\|_i$.
(3) The need for this estimate is explained following Corollary 6.5.

*Proof of Lemma 7.7.* Fix $p$ such that $\frac{2\dim(\mathfrak{L}_x)}{p} < 1$. For a choice of $0 < C < 1$ and any choice of $\varsigma_1 > 0$ and $F > 0$, choose a neighborhood $U$ in $\mathrm{Diff}^K(X, \mathfrak{F})$ and a function $h$ in $\mathcal{U}(\Gamma)$ satisfying Proposition 5.4. Then for any $(U, C^k)$-foliated perturbation $\rho'$ of $\rho$, we have that $\rho'_P(h)^n f$ converges pointwise almost everywhere to $\rho'_P$ invariant function $f'$ with respect to $M\mu$ and that:

$$\|\rho'_P(h)^{n+1} f - \rho'_P(h)^n f\|_{p,k,M\mu} \leq C^n F.$$

Let $0 < D = \sqrt{C} < 1$ and applying Lemma 7.5 shows that there exists a set $S_n$ such that $\mu(S_n^c) < D^n \delta$ and for every point $x \in S_n$, we have that

$$\|(\rho'_P(h)^{n+1} f - \rho'_P(h)^n f)|_{B_{\mathfrak{F}}(x,r)}\|_{k-1} \leq D^n A F \tag{6}$$

where $A > 0$ is an absolute constant depending only on $(X, \mathfrak{F}, g_{\mathfrak{F}})$.

We will show that, possibly after shrinking $U$, $\rho'_P(h)^n f$ satisfies

$$\|(\rho'_P(h)^{n+1}f-\rho'_P(h)^n f)|_{\pi_1{}^{-1}(B_{\mathfrak{F}}(x,r))}\|_l \leq D'^n P(nF_l)F \tag{7}$$

for $n > j(x)$ where $j$ is an integer valued measurable function on $X$, where $P$ is a fixed polynomial, and $F_l > 0$ and $0 < D' = D'(D,l,h) < 1$. This estimate immediately implies that $\rho'_P(h)^n f|_{\pi_1{}^{-1}(B_{\mathfrak{F}}(x,r))}$ converges in $C^l(\pi_1{}^{-1}(B_{\mathfrak{F}}(x,r)))$ which suffices to complete the proof.

We let $b = 2l-k+1$ and can now define $F_l = \sup_{\mathrm{supp}(h)}\|\rho'(\gamma)\|_b$. We shrink $U$ so that $\|\rho'(\gamma)\|_1^b D < 1$ for every $\gamma \in \mathrm{supp}(h)$, let $D_h = \sup_{\mathrm{supp}(h)}\|\rho'(\gamma)\|_1$. We also fix the constant $D' = D'(l,h,D)$ such that $\sqrt{D_h^b D} < D' < 1$. Letting $f_n = \rho'_P(h)^n(\rho'_P(h)f - f)$ and $f_n^x = f_n|_{B_{\mathfrak{F}}(x,r)}$, Lemma 6.3 implies that

$$\|f_n^x\|_l \leq B\|f_n^x\|_{k-1}^{\frac{1}{2}}\|f_n^x\|_b^{\frac{1}{2}} \tag{8}$$

for $B$ depending only on $X, \mathfrak{F}, b$.

We now form the product $\Omega \times X$ with measure $\mu \times \lambda$ and transformation $T_{\rho'}$ as defined in the paragraph immediately preceding the proof. Define subsets $\tilde{S}_n = \Omega \times S_n$. We now define sets $\bar{S}_j = \cap_{i=j+1}^{\infty} T^{-i}\tilde{S}_i$. This is the set of $(\omega,x) \in X$ such that $T_{\rho'}^i(\omega,x) \in \tilde{S}_i$ for all $i > j$. The Borel-Cantelli lemma implies that $\cup \bar{S}_j$ has full measure in $\Omega \times X$. The function $j$ will be defined so that $j(x)$ is the smallest integer such that $x \in \pi_X(\bar{S}_j)$ and we will prove inequality (7) by fixing $j$ and assuming $x \in \pi_X(\bar{S}_j)$.

Applying inequality (6) to any point in $\pi_X{}^{-1}(x) \cap \bar{S}_j$ implies that

$$\|f_n^x\|_{k-1} \leq D^n AF \tag{9}$$

for every $x$ with $x \in \pi_X(\bar{S}_j)$ whenever $n > j$. It remains to find a bound on $\|f_n^x\|_b$. Noting that $f_n = \rho'_P(h)^n(\rho'_P(h)f - f)$ Corollary 6.5 implies that

$$\|f_n^x\|_b \leq D_h^{nb} P(nF_l) \tag{10}$$

where $P$ is a constant multiple of the polynomial occurring in Corollary 6.5. Inequality (7) is now immediate from inequalities (8), (9) and (10) and the definition of $D'$. □

## Appendix A. "Good spaces" for continuous limit actions

The purpose of this appendix is to show how to adapt the argument given in subsection 3.3 to prove the general cases of Proposition 3.12 and Proposition 3.13. The proof of Proposition 3.12 is completed in the first two subsections and the third subsection ends with the proof of Proposition 3.13. More generally, this appendix contains a series

of remarks concerning the category of spaces and actions which admit "good" limit actions, as well as characterizations of certain of these spaces.

A.1. **Triangles and convexity of continuous subactions.** In this subsection we outline a proof that, under the hypotheses of Proposition 3.12 $\rho = \omega\text{-lim}\,\rho_n$ is continuous on an affine subspace. To see this it suffices to study sequences of triples $A_n, B_n, C_n \in \mathcal{H}_n$ such that $C_n = tA_n + (1-t)B_n$ and show that equicontinuity at $A_n$ and $B_n$ implies equicontinuity at $C_n$. This follows from the fact that almost isometries are almost affine, i.e. that the image of a convex combination of points under an almost isometry is close to the same convex combination of the images of the points. We state this fact precisely only for globally defined actions, though a more complicated analogue is clearly true for partially defined actions.

**Lemma A.1.** *For every $\eta > 0, t_0 > 0$ and $R > 0$ there exists $\varepsilon > 0$ such that if $f$ is an $\varepsilon$-almost isometry of a Hilbert space $\mathcal{H}$ and $A, B \in \mathcal{H}$ with $d(A,B) < R$ and $t < t_0$ then*

$$d(f(tA + (1-t)B), tf(A) + (1-t)f(B)) < \eta$$

As the lemma is easily proved from elementary facts concerning stability of triples of collinear points in a Euclidean space, we only indicate what is needed for the proof. Take three collinear points $A, B, C$ and three arbitrary points $A', B', C'$ such that $d(A,B) \simeq d(A',B'), d(A,C) \simeq d(A',C')$ and $d(B,C) \simeq d(B',C')$. Then the triangles $\Delta(ABC)$ and $\Delta(A'B'C')$ are almost congruent. More precisely, if we move $\Delta(A'B'C')$ by an isometry so that $A = A'$ and so that $B'$ is as close as possible to $B$, then $C$ will be close to $C'$. We leave precise quantification of this fact to the interested reader. As Lemma A.1 uses only this fact about triangles, it is clear that the lemma is true much more generally. For example, the lemma holds for any $L^p$-type space where $1 < p < \infty$, as well as for $CAT(0)$ spaces. The lemma, and therefore the first conclusion of Proposition 3.12, should hold for any geodesic metric space which is uniformly convex in any reasonable sense, see below or [KM] for possible definitions. If the space does not admit a linear structure, one needs to interpret affine subspaces and affine combinations in terms of the geodesic structure.

A.2. **Barycenters, uniform convexity and almost isometries.** In this subsection we indicate the proof of the remaining conclusion of Proposition 3.12. Given a metric space $Y$, let $\mathcal{P}(Y)$ be the set

of regular, Borel probability measures on $Y$. Given $\mu\in\mathcal{P}(Y)$, we define $f_\mu(x) = \int_Y d(y,x)^2 d\mu(y)$. If $f_\mu$ attains a global minimum at a unique point, we call that point the *barycenter* of the measure, and we denote by $b : \mathcal{P}(Y)\to Y$ the map taking a measure to its barycenter (when it exists). For any point $x_0$ in a Hilbert space $\mathcal{H}$, the function $f_{x_0} = d(x_0,x)^2$ has the property that it's restriction to any geodesic has second derivative 2 at every point. By definition this property is inherited by $f_\mu$ for any measure $\mu$. This implies that $f_\mu$ has at most one minimum and easily implies that the barycenter is defined at least when the support of $\mu$ is compact. The barycenter is not defined for $\mu$ with non-compact support as can be seen by taking an atomic measure supported on an infinite sequence of points $\{x_n\}$ which go to infinity much faster than $\mu(x_n)$ goes to zero. More generally, barycenters will exist for measures which decay fast enough at infinity. We leave the precise formulation to the reader.

The relevance of this discussion for subsection 3.3 follows from the fact that for Hilbert spaces $b(\mu) = \int_{\mathcal{H}} v d\mu(v)$. This is easily seen by showing that $\int_{\mathcal{H}} v d\mu(v)$ is a critical point for $f_\mu$. Combined with our observation on the second derivative of $f_\mu$ along any geodesic this implies:

**Lemma A.2.** *For every Hilbert space $\mathcal{H}$ and every compactly supported $\mu\in\mathcal{P}(Y)$, the function $f_\mu$ has a unique global minimum $m_\mu$ at a point $y_\mu = \int_{\mathcal{H}} v d\mu(v)$. Furthermore, for every $\varepsilon > 0$ and any compactly supported probability measure $\mu$ on any Hilbert space $\mathcal{H}$ the set of points where $f_\mu(x) < m_\mu + \varepsilon$ is contained in $B(b(\mu), \sqrt{\varepsilon})$.*

It is immediate from the definition that $b$ is $\mathrm{Isom}(\mathcal{H})$ equivariant. We now describe the behavior of $b$ under $\varepsilon$-almost isometries.

**Lemma A.3.** *For every $D, \varepsilon > 0$ there is an $\eta > 0$ such that if $\mathcal{H}$ is a Hilbert space, $y_0\in\mathcal{H}$ is a basepoint and $\mu\in\mathcal{P}(\mathcal{H})$ with $m_\mu < D$ and $\mathrm{supp}(\mu)\in B(y_0,R)$ and $g$ is a $\eta$-almost isometry from $B(y_0,R)$ to $Y$, then*

$$d(g(b(\mu)), b(g_*\mu)) < \varepsilon.$$

*Proof.* Since $g$ is an $\eta$-almost isometry, we know that

$$(1-\eta)d(x,y)\leq d(g(x),g(y))\leq(1+\eta)d(x,y)$$

for every $x,y\in B(y_0,R)$. Since $g_*\mu(S) = \mu(g^{-1}(S))$, squaring and integrating implies that $(1-\eta)^2 f_\mu(y)\leq f_{g_*\mu}(g(y))\leq(1+\eta)^2 f_\mu(y)$. In particular $(1-\eta)^2 m_\mu\leq m_{g_*\mu}\leq(1+\eta)^2 m_\mu$ and therefore $f_\mu(b(g_*\mu))\leq(1+\eta)^4 m_\mu$. Combined with Lemma A.2 this implies that

$$d(g(b(\mu)), b(g_*\mu)) < \sqrt{((1+\eta)^4-1)D}.$$

□

To complete the proof of Proposition 3.12 it suffices to show that $\{y_n\}$ is in $C$ whenever $y_n = \rho_n(f)z_n$ for $z_n \in \tilde{X}$ and $f \in \mathcal{U}(\Gamma)$. Letting $\mu_n$ be the push-forward of $f d\mu_\Gamma$ under $\rho_n^{z_n}$ we need to show that for every $\varepsilon > 0$, there exists a neighborhood of the identity $U$ in $\Gamma$ such that $d_n(\gamma y_n', y_n') < \varepsilon$ for every $\gamma \in \Gamma$. Note that $d_n(\gamma y_n', y_n') \le d_n(\gamma y_n', b(\gamma\mu_n)) + d(b(\gamma\mu_n), b(\mu_n))$. The first term can be made arbitrary small by Lemma A.3 since $\omega\text{-}\lim \varepsilon_n = 0$. Bounding the second term follows as in the proof of the affine case of Proposition 3.13.

**Remarks:**

(1) We can define the *lower second derivative* of a function $f : \mathbb{R} \to \mathbb{R}$ by
$$\underline{f}''(x) = \liminf_{h\to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}.$$
For any CAT(0) space $Y$ and any point $y_0$, it is easy to show that the restriction of $f_{y_0}(y) = d(y, y_0)^2$ to any geodesic satisfies $\underline{f}''_{y_0} \ge 2$. Only slightly more difficult is showing that this property characterizes CAT(0) spaces. A similar remark is made in [Gr2].

(2) A harder exercise is to show that if for every point $y_0 \in Y$ and every geodesic $c$ in $Y$, we have $\underline{f}''_{y_0} = 2$ on $c$, then $Y$ is Hilbert space.

(3) An analog of Lemma A.3, and therefore Proposition 3.12, is true for more general spaces $X$ in place of the Hilbert space $\mathcal{H}$, provided we define $\rho(h)x = b(\nu)$ where $\nu$ is push-forward of $h\mu_\Gamma$ under the orbit map $\rho(\Gamma)x \to X$. In particular, if all spaces acted upon are CAT(0) spaces or are $L^{p_n}(Y,\nu)$ for $(Y,\nu)$ a standard measure space and $1 < \omega\text{-}\lim p_n < \infty$. More generally, this will be true for any uniformly convex metric space in the sense of say [KM].

A.3. **Convolutions and linear structure.** We will now proceed to prove Proposition 3.13 from the following lemma and Lemma A.3.

**Lemma A.4.** *Let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$ be finite subsets of a Hilbert space $\mathcal{H}$ such that $d(x_i, y_i) < \eta$ for all $i$. Then for any coefficients $a_1, \ldots, a_n \in \mathbb{R}$, we have $d(\sum_{i=1}^n a_i x_i, \sum_{i=1}^n a_i y_i) < (\sum_{i=1}^n |a_i|)\eta$.*

*Proof.* All statements are easy consequences of the triangle inequality for the norm on the Hilbert space and the fact that $d(u,v) = \|u - v\|$. □

*Proof of Proposition 3.13.* Since atomic measures with finite support are dense in $\mathcal{P}(\Gamma)$ we assume that $\mu = \sum_i a_i\delta_{\gamma_i}$ where the $g_i \in G$ and $a_i$ are positive reals. First we note that $\rho(\mu * \lambda)x = \sum_i a_i b(g_i\rho^x_*\lambda)$ and $\rho(\lambda)x = b(\rho^x_*\lambda)$ where $\rho^x : G \to G\cdot x$ is the orbit map. Now $\rho(\mu)\rho(\lambda)x = \sum_i a_i g_i b(\rho^x_*\lambda)$. By Lemma A.3 applied to each $g_i$ and the measure $\rho^x_*\lambda$ we have that $d(g_i b(\rho^x_*\lambda), b(g_i\rho^x_*\lambda)) \leq \eta$. The Proposition now follows from Lemma A.4. □

Unlike Proposition 3.12, Proposition 3.13 holds in much less generality, since it depends on the affine structure of $X_n$ and the equation $b(\mu) = \int_{X_n} v d\mu$. In fact, to prove more general variants of our results it is probably best to simply define $\rho(h)x = b(\rho^x_* h\mu_\Gamma)$ and work with this averaging operator instead of the linear one.

## Appendix B. Estimates on Compositions

This appendix contains a proof of Lemma 6.4. Given the definitions, it suffices to prove the Lemma for $\phi \in \mathrm{Diff}^k(X, \mathfrak{F})$. We deduce this from some elementary facts about block upper triangular matrix.

Given a number $N$, we consider $N \times N$ matrices which are *block upper triangular*. By this we mean that there are number $i_1, \ldots i_n$ such that $\sum_{l=1}^n i_l = N$ and the matrices $M$ have $i_l \times i_l$ blocks, which we denote $A_l$, along the diagonal, are zero below these blocks, and have arbitrary entries above them. We call such $M$ *block upper triangular of type* $i_1, \cdots, i_n$. We define a norm on matrices by taking the maximum of the matrix coefficients. It is easy to see that this is equivalent to the operator norm.

**Lemma B.1.** *Let $M_1, \ldots, M_j$ be block upper triangular matrices of type $i_1, \ldots, i_n$. Assume that $\|A_l\| < C_1$ for all $l$ and that all other entries of each $M_k$ are bounded by $C_2$. Then there exists a polynomial $Q$ depending on the type of the $M_l$ such that*

$$\|M_1 M_2 \cdots M_j\| \leq C_1^j Q(jC_2).$$

*Proof.* It is easy to see that the diagonal blocks of $M = M_1 M_2 \cdots M_j$ satisfy this bound, and in fact are less than $C_1^j$. For any coefficient of the product outside of the diagonal blocks, we can write the matrix coefficients of $M$ as:

$$M_{\alpha,\beta} = \sum_{\alpha=\eta_0 \leq \eta_1 \cdots \leq \eta_{j-1} \leq \beta = \eta_j} (M_1)_{\alpha,\eta_1}(M_2)_{\eta_1,\eta_2} \ldots (M_j)_{\eta_{j-1},\beta}.$$

It is easy to see that at most $N$ of the $(M_l)_{\eta_{l-1},\eta_l}$ can be outside the diagonal blocks of $M_l$ (or even off the diagonal), since each entry of this form has $\eta_l > \eta_{l-1}$. The number of choices of such sequences is

$\binom{j}{N}$ which is a polynomial $Q$ in $j$ of degree $N$. The norm of $M_{\alpha\beta}$ is then bounded by $C_1^j Q(jC_2)$ as desired. □

*Proof of Lemma 6.4.* We use the fact that $J^k(X,\mathfrak{F})\simeq\bigoplus_{j=0}^k(S^j(T\mathfrak{F}^*))$. Given $\phi\in\mathrm{Diff}^k(X,\mathfrak{F})$ and $x\in X$, we can write $j^k(\phi)(x)$ with respect to bases of $J^k(X,\mathfrak{F})$ at $x$ and $\phi(x)$ which respects this splitting. Then it is clear that $j^k(\phi)(x)$ is block upper triangular, where the diagonal blocks are of the form $S^j(D\phi)(x) = S^j(j^1(\phi))(x)$ where $j = 0,\dots,k$. Therefore the norm of the blocks is bounded by $N_1^k$, and Lemma 6.4 is an immediate consequence of Lemma B.1. □

## Appendix C. Locally compact groups and free topological groups

In this appendix we sketch a proof of the following Proposition. We believe this Proposition to be well-known, and experts we consulted all provided proofs more or less along the following lines, but none could provide a reference.

**Proposition C.1.** *Let $\Gamma$ be a locally compact, $\sigma$-compact topological group and $K$ a compact generating set containing a neighborhood of the identity. Then the group $\Gamma'$, generated by $K$ and satisfying all the relations of $\Gamma$ of the form $xy = z$ where $x, y$ and $z$ are in $K$ can be given a topology as a locally compact, $\sigma$-compact group.*

We first note the following lemma.

**Lemma C.2.** *Let $\Gamma$ be a locally compact, $\sigma$-compact group. Then $\Gamma$ admits a left invariant metric $d_L$ which defines the topology.*

*Proof.* It is well-known that $\Gamma$ is completely regular and therefore metrizable and so we can choose a metric $d$ on $\Gamma$, which is not necessarily left invariant. To find a left invariant metric we choose a continuous function $f$ supported on a compact neighborhood $C$ of the identity such that $f$ is continuous, $f = 1 - d(x, y)$ on a smaller neighborhood of the identity and $f\equiv 0$ outside $C$. We then define a map from $\Gamma$ to continuous function on $\Gamma$ by $\gamma\to\gamma\cdot f$ and define a function $d_L$ on $\Gamma\times\Gamma$ by $d_L(\gamma_1, \gamma_2) = \|\gamma_1\cdot f - \gamma_2\cdot f\|_{C^0}$. This is clearly left invariant and it is easy to check that $d_L$ is a metric and defines the same topology as $d$ on $\Gamma$. □

It is clear from the construction of $d_L$ that we can normalize so that the ball of radius one is contained in $K$.

To prove the proposition, we need to define a topology on $\Gamma'$. The group $\Gamma'$ as a group is the quotient of the free group $F(K)$ and we

call the projection map $\pi$. We can define a norm on $F(K)$ by letting $\|k_1k_2\ldots k_n\|_F = \sum_{i=1}^n d_L(e_\Gamma, k_i)$ and then define

$$\|\gamma\|_{\Gamma'} = \inf_{\{w\in F|\pi(w)=\gamma'\}} \|w\|_F.$$

The fact that $\|\cdot\|$ is a norm is straightforward and we define a topology on $\Gamma'$ by taking a system of neighborhoods of a point $x$ to be sets of the form

$$\{y|\,\|x^{-1}y\|_{\Gamma'} < \frac{1}{n}\}.$$

It remains to check that this defines a topology on $\Gamma'$ that makes $\Gamma'$ a topological group, and we will indicate a proof of this below, though it also follows easily from results in [Ma]. For more details on norms, topologizing topological groups via norms and a construction of a topology on $F(K)$ which makes $\pi$ continuous, see [Ma]. We now note an essentially trivial lemma. We leave the proof to the reader.

**Lemma C.3.** *For any $\gamma$ in $\Gamma'$ with $\|\gamma\|_{\Gamma'} < 1$, we can write $\gamma = \pi(k)$ for some $k\in K$ such that $\|\gamma\|_{\Gamma'} = d(e_\Gamma, k)$.*

Let $U$ be the set of words in the free group with $\|w\|_F < 1$. As immediate consequences of the lemma we have:

(1) the map from $\Gamma'\to\Gamma$ is a homeomorphism on the set $\pi(U)$ in $\Gamma'$
(2) the set $\pi(U)$ contains a neighborhood of the identity in $\Gamma'$.

It only remains to check that the topology we have defined on $\Gamma'$ makes $\Gamma'$ a topological group. To see this one merely needs to check that the topology is invariant under conjugation in some neighborhood of the identity. To check this, it suffices to check it for conjugation by elements of $K$, but there it is more or less obvious, as the action of $K$ by conjugation on a small enough neighborhood of the identity contained in $\Gamma'$ is now easily seen to be conjugate by a homeomorphism to the action of $K$ by conjugation on a small neighborhood of the identity in $\Gamma$.

## Appendix D. Historical remarks, relations to other work, and further generalizations

This appendix attempts to clarify the relationship of our work to the work of others and also contains some remarks that may be useful for future generalizations of our results.

D.1. **Proofs of Theorem 1.1 and 2.11 and KAM theory.** This subsection first discusses the failed proof mentioned in the introduction to this paper, and then goes one to compare that failed proof, the current successful one, and the KAM method.

We recall our original approach to proving Theorem 1.1. Given an isometric action $\rho$ of $\Gamma$ on a compact manifold $X$ and a perturbation $\rho'$ of $\rho$, a conjugacy is a diffeomorphism $f : X \to X$ such that $\rho(\gamma)\circ f = f\circ\rho'(\gamma)$ for all $\gamma$ in $\Gamma$. Rearranging, the conjugacy is a fixed point for the $\Gamma$ action on the group $\mathrm{Diff}^k(X)$ of diffeomorphisms of $X$ defined by $f \to \rho(\gamma)\circ f\circ\rho'(\gamma)^{-1}$. Ideally we would parameterize diffeomorphisms of $X$ locally as a Hilbert space and then use Theorem 2.3 below to find a fixed point or conjugacy.

We briefly describe an approach to this parametrization and the difficulty encountered. Let $\mathrm{Vect}^k(X)$ be the set of vector fields on $X$ and $C^k(X, X)$ be the set of $C^k$ maps from $X$ to $X$. Given a Riemannian metric on $X$ there is a natural exponential map $\mathrm{Exp} : \mathrm{Vect}^k(X) \to C^k(X, X)$ defined by taking a vector $V$ to the time one map of the geodesic flow along $V$ and projecting back to $X$, i.e. by $V \to (x \to \exp_x V_x)$. If $\rho = \rho'$ and we define Exp using the $\rho$ invariant metric, we have a natural action of $\Gamma$ on $\mathrm{Vect}^k(X)$ such that Exp is equivariant. As will be shown in section 4 it is fairly straightforward to complete $\mathrm{Vect}^k(X)$ with respect to a Sobolev metric in such a way that the completion $\mathrm{Vect}^{2,k}(X)$ is a Hilbert space on which the $\Gamma$ action defined by $\rho$ is isometric. We had hoped to show that if $\rho'$ is close enough to $\rho$ then we would have a partially defined $\Gamma$ action on $\mathrm{Vect}^{2,k}(X)$ that was by $\varepsilon$-almost isometries. We would then apply Theorem 2.3 to this partially defined action to find a fixed vector field $V$, and $\mathrm{Exp}(V)$ would be the desired conjugacy. While it is possible to construct a partially defined action on a ball in $\mathrm{Vect}^k(X)$, we were unable to show that the action is $\varepsilon$-almost isometric if one considers a metric on $\mathrm{Vect}^{2,k}(X)$ with $k > 1$. This is important, since to show that $V$ and $\mathrm{Exp}(V)$ are smooth, and that $\mathrm{Exp}(V)$ is invertible, one needs to use the Sobolev embedding theorems, which require a loss of derivatives proportional to the dimension of $X$. This method fails even if we could use an $L^p$ type Sobolev space $\mathrm{Vect}^{p,k}$ and Corollary 2.8, since at most these results will produce a continuous invariant vector field $V$, and it is not clear that $\mathrm{Exp}(V)$ is even a homeomorphism. We remark that we cannot use Corollary 2.8 since the action on $\mathrm{Vect}^k(X)$ is not linear.

It is worth noting that this is different than the difficulty with loss of derivatives usually encountered by $KAM$ type methods. Here the problem is that no matter what topology we assume $\rho$ and $\rho'$ are close in, we cannot prove that the action we define on vector fields satisfies *any estimate of any kind* on higher derivatives. In KAM the typical problem is that estimates for the solutions to the linearized equation

are only “uniformly good” for low order, but one has some a priori estimate at higher order.

This difficulty arises from the fact that the method suggested does not actually involve linearizing the action on diffeomorphisms. What we had hoped to do was to use the parametrization mentioned to provide a linear structure in which the action constructed was “almost isometric”, and therefore “close enough” to linear, so as to be able to apply Theorem 2.5. The problem is that we can only do this in function spaces where this yields no meaningful results.

To resolve this difficulty, we linearize the problem, and in fact give two different linearizations. Our linearizations are not very similar to the $KAM$ linearization, and their utility depends heavily on our results concerning groups with property $(T)$. More or less our method takes advantage of the fact that, for groups with property $(T)$, contracting properties of certain operators are preserved under small perturbations for actions on a wide variety of uniformly convex Banach spaces. The disadvantage of our method is that to obtain such a perturbation, we need to only consider Banach spaces whose definition involve only finitely many derivatives. For a long time, this left a $C^{\infty,\infty}$ result out of reach. Our proof of the $C^{\infty}$ case was inspired by a study of the $KAM$ method and particularly of the paper [DK], but the only concrete similarities to $KAM$ arguments is the use of an iteration and the types of estimates used.

From the point of view of $KAM$ theory it is surprising that we need the estimates from Lemma 6.4 given the strong contracting properties of the averaging operators we consider. The need for these estimates is explained following Corollary 6.5 and Proposition 4.2. It is possible to give a proof of the $C^{\infty,\infty}$ case of Theorem 1.1 without using these estimates. This has been done very recently by the first author using Hamilton's implicit function theorem and an approach similar to Weil's work on local rigidity of lattices in Lie groups [F]. This approach has applications to local rigidity of isometric actions for some groups that do not have property $(T)$, but is unlikely to yield a result in the generality of Theorem 2.11. The proof uses many facts concerning harmonic analysis on compact manifolds that are unknown, unlikely to be true, or known to be false in the context of general compact foliated spaces.

D.2. **Further fixed point properties and relations to the work of M.Gromov.** Examination of the proof of Theorem 2.5 shows that one can state more general variants of the theorems discussed here. The limiting procedure applied in the proof is quite flexible, and allows

one to limit over almost any set of parameters. To some extent this is illustrated in the proofs of the results from subsection 2.2. In another direction, one can replace $\mathcal{H}$ by a non-positively curved space that is "$\varepsilon$-almost flat". By this one should mean anything that implies that, given a sequence of "$\varepsilon$-almost flat" spaces with $\varepsilon$ tending to zero, the limit space constructed by the method of subsection 3.2 is a Hilbert space. To actually prove this variant, we need to define the operator $\rho(h)$ for actions (or partially defined actions) of $\Gamma$ on spaces of non-positive curvature. A method for doing this is described in subsection A.2 of the appendix. Theorem 2.3 can be generalized even further to "$\varepsilon$-almost flat" spaces which are not non-positively curved see the discussion related to Lemma D.1 below. For discrete groups, these more general assertions are easy exercises from the proofs in section 3 below. For non-discrete groups, the issue of finding a continuous subaction of the limit action constructed in subsection 3.2 can present non-trivial difficulties or require additional assumptions.

In [Gr2], Gromov proves that certain "random" infinite, discrete groups have a fixed point property that is stronger than property $(T)$. He proves that these groups have fixed points for any isometric action on any finite or infinite dimensional "regular" non-positively curved space. After having completed an earlier draft of this paper, we discovered that the ideas in [Gr2] have many points in common with ours. In particular, in section 3.13B, Gromov outlines a proof of a special case of Theorem 1.6, for a certain class of "random" infinite, discrete groups with property $(T)$ and for affine actions. This is a class of groups whose Cayley graphs "contain" a family of expander graphs as subsets. By a graph being contained in the Cayley graph, we mean that the Cayley graph contains an embedded copy of the graph. Actually, the Cayley graphs of Gromov's groups only contain "most" of the relations that would arise from containing the collection of expander graphs in a sense made precise in [Gr2]. By a family of expander graphs we mean a collection of $(n, k, c)$ expanders with $k$ and $c > 0$ fixed and $n$ going to infinity. Although one can build a family of expander graphs of this kind as a series of quotients of any residually finite group with property $(T)$, it is far from clear that one can realize a family of expanders as subsets of the Cayley graph for an arbitrary discrete group with property $(T)$, even in Gromov's probabilistic sense.

More generally, a central philosophy of [Gr2] is that if a collection of spaces $\mathcal{C}$ is "closed under scaling limits" then for a group $\Gamma$ to have almost fixed points (i.e. sequences of points with $K$-displacement converging to zero) for all isometric actions on spaces in $\mathcal{C}$ is equivalent to having fixed points for all such actions. From this point of view the

emphasis of our results on groups with property $(T)$ is on extending the fixed point property to (partially defined) actions that are close enough to being isometric on spaces that are "close enough" to $\mathcal{C}$. Finally, we note that it should be possible to prove a common generalization, and show that Gromov's groups have fixed points for partially defined $\varepsilon$-almost isometric actions on "regular" non-positively curved spaces.

A primary technical difference between our work and Gromov's is the functional used. Where we use the $K$ displacement, Gromov uses a $K$ energy. Despite this, various variants of Proposition 3.1, for discrete groups with property $(T)$, permeates section 3 of [Gr2], see particularly $3.8-3.13$. The precise formulations given there are somewhat more complicated because they are phrased in terms of energy rather than displacement. For a gentler presentation of some of the ideas in [Gr2], see the commentaries [Si] and section 6 of [Gh].

**On the use of ultrafilters:** It is possible to construct the limit isometric action of $\rho$ on a Hilbert space $\mathcal{H}$ "by hand" without using ultrafilters, at least when $\Gamma$ is discrete. To do this, one chooses an explicit isometric identification of the orbits $\rho_n(\Gamma)x_n$ with subsets of a fixed Hilbert space $\mathcal{H}$, always identifying $x_n$ with 0. By passing to a subsequence were $\{\rho_n(\gamma)x_n\}$ converges for every $\gamma$, we can obtain an isometric action of $\Gamma$ on a countable set in $\mathcal{H}$ that extends to an action on a closed linear subspace of $\mathcal{H}$. Verifying this and then obtaining the contradiction between the properties of the $\Gamma$ action on $\mathcal{H}$ and the $\Gamma$ action on $\mathcal{H}_n$ is considerably more involved than the proof above, though the argument does not use much more than simple linear algebra and geometry. The argument is similar to the proof of Proposition 3.13 and uses some of the same lemmas. It does not seem possible to carry out arguments of this type in the generality of subsection 3.2.

**Fixed points without iterative method:** If one is more interested in Theorem 2.3 than Theorem 2.5, it is possible to provide an independent proof of that theorem along the same lines. This may be useful for generalizations to spaces where the operators $\rho(h)$ are either not defined or not well-behaved. To do so one needs to produce a Cauchy sequence of points with smaller and smaller displacement for all partially defined actions which are "close enough" to being isometric actions. The following lemma, stated by the second author in Jerusalem in 1997, suffices:

**Lemma D.1.** *Let $\Gamma$ be a group with property $(T)$ and fix a compact generating set $K$. Given $\delta_0 > 0$ there exist $\varepsilon > 0, r = r(\delta_0) > 0$, and positive integers $s$ and $M$, such that for any Hilbert space $\mathcal{H}$, any $\delta < \delta_0$*

*and any* $x \in \mathcal{H}$, *and any continuous* $(r, s, \varepsilon, \delta, K)$*-action of* $\Gamma$ *on* $\mathcal{H}$ *one can find a point* $y$ *such that:*

(1) $d(x,y) \leq M \operatorname{disp}_K(x)$ *and,*
(2) $\operatorname{disp}_K(y) \leq \frac{1}{2} \operatorname{disp}_K(x)$.

To prove Lemma D.1 one argues by contradiction as in the proof of Theorem 2.5. The contradiction follows since if $\{x_n\}$ is our sequence of basepoints, then there is a fixed point $y_\omega$ in the limit action with $d(x_\omega, y_\omega)$ less than $M$ times the $K$ displacement of $x_\omega$, where $M > 0$ is a constant depending only on $\Gamma$ and $K$. (This fact for isometric actions on Hilbert spaces is, for example, an easy corollary of Proposition 3.1.) To prove Theorem 2.3 one then argues as in the proof that Theorem 2.1 implies Theorem 1.6.

Lemma D.1 suffices to prove Theorems 1.6 and 2.3, but these results do not suffice for our applications. In particular, we need the precise iterative method of finding fixed points:

(1) to obtain optimal regularity by finding estimates in $L^p$ type Sobolev spaces for large $p$,
(2) to control the non-uniformities that arise in applying the Sobolev embedding theorems on foliated spaces and,
(3) to be able to use the estimates of Section 6 to obtain $C^{\infty,\infty}$ results.

## References


[Ab] Abels, Herbert *Finite presentability of $S$-arithmetic groups. Compact presentability of solvable groups.* Lecture Notes in Mathematics, 1261. Springer-Verlag, Berlin, 1987.

[BFGM] U.Bader, A.Furman, T.Gelander, N.Monod, Property $(T)$ and actions on $L^p$ spaces, preprint.

[BL] Y.Benjamini, J.Lindenstrauss, *Geometric Non-linear Functional Analysis: Volume 1*, Colloquium Publications, AMS, 2000.

[Be] E.J.Benveniste, Rigidity of isometric lattice actions on compact Riemannian manifolds, *GAFA* 10 (2000) 516-542.

[BTG] N. Bourbaki, Éléments de mathématique. Topologie générale. Chapitres 1 4. Hermann, Paris, 1971.

[CC] A.Candel and L.Conlon, *Foliations I*, American Mathematical Society Graduate Studies in Mathematics, Volume 23, Providence, 2000.

[doC] M.P. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[De] Delorme, Patrick 1-cohomologie des reprsentations unitaires des groupes de Lie semi-simples et rsolubles. Produits tensoriels continus de reprsentations. (French)*Bull. Soc. Math. France* 105 (1977), no. 3, 281–336.

[D] Dieudonne, *Elements d'Analyse*, vol. 4, Gauthiers-Villars, Paris, 1971.

[DK] D. Dolgopyat and R. Krikorian, On simultaneous linearization of diffeomorphisms of the sphere, preprint.

[F] D.Fisher, Harmonic analysis, Hamilton's implicit function theorem and local rigidity, in preparation.

[FM1] D.Fisher and G.A.Margulis, Local rigidity for cocycles,in *Surv. Diff. Geom. Vol VIII*, refereed volume in honor of Calabi, Lawson, Siu and Uhlenbeck , editor: S.T. Yau, 45 pages, 2003.

[FM2] D.Fisher and G.A.Margulis, Local rigidity of affine actions of higher rank groups and lattices, preprint.

[Gh] E.Ghys, Groupes aléatoire [d'aprés Misha Gromov], to appear *Sem. Bourbaki.*

[Gr1] M. Gromov, Asymptotic invariants of infinite groups.*Geometric group theory, Vol. 2 (Sussex, 1991)*, 1–295, London Math. Soc. Lecture Note Ser., 182, Cambridge Univ. Press, Cambridge, 1993.

[Gr2] M.Gromov, Random walk in random groups, *GAFA* 13 (2003) vol 1, 73-146.

[Gu] Guichardet, Alain Sur la cohomologie des groupes topologiques. II. (French) *Bull. Sci. Math.* (2) 96 (1972), 305–332.

[HV] P. de la Harpe, A.Valette, *La propriete (T) de Kazhdan pour les groupes localement compacts*, Asterisque 175, Soc. Math. de France, Paris, 1989.

[He] Heinrich, S, Ultraproducts in Banach space theory, *J. Reine Angew. Math.* 313 (1980) 72-104.

[H] Hirsch, Morris W. *Differential topology.* Graduate Texts in Mathematics, No. 33. Springer-Verlag, New York-Heidelberg, 1976.

[Ho] Hörmander, Lars The boundary problems of physical geodesy. *Arch. Rational Mech. Anal.* 62 (1976), no. 1, 1–52.

[K] D. Kazhdan, On the connection of the dual space of a group with the structure of its closed subgroups. (Russian) *Funct. Anal. Appl.* 1 1967 71–74.

[KM] A.Karlsson and G.A.Margulis. A multiplicative ergodic theorem and nonpositively curved spaces. *Comm. Math. Phys.* 208 (1999), no. 1, 107–123.

[KN] S.Kobayashi and K.Nomizu, *Foundations of Differential Geometry, Vol. I*, Wiley-Interscience Publications, New York, 1963.

[M] G.A. Margulis, *Discrete subgroups of semisimple Lie groups*, Springer-Verlag, New York, 1991.

[Ma] Markov, A. A. Three papers on topological groups: I. On the existence of periodic connected topological groups. II. On free topological groups. III. On unconditionally closed sets. *Amer. Math. Soc. Translation*, (1950). no. 30.

[MU] S.Mazur, S.Ulam, Sur les transformations isomtriques d'espaces vectoriels normes, *C.R. Acad. Sci. Paris* 194 (1932) 946-948.

[MS] C.C. Moore and C. Schochet, *Global analysis on foliated spaces.* With appendices by S. Hurder, Moore, Schochet and Robert J. Zimmer. Mathematical Sciences Research Institute Publications, 9. Springer-Verlag, New York, 1988.

[Mo] Mostow, G. D. Equivariant embeddings in Euclidean space. *Ann. of Math.* (2) 65 (1957), 432–446.

[P1] Palais, R.S., Differential operators on vector bundles, *Seminar on the Atiyah-Singer index theorem*, Annals of Math. Studies 57, Princeton University Press, Princeton, 1965.

[P2] Palais, Richard S. Imbedding of compact, differentiable transformation groups in orthogonal representations. *J. Math. Mech.* 6 (1957), 673–678.

[S] Y.Shalom, Rigidity of commensurators and irreducible lattices. *Invent. Math.* 141 (2000), no. 1, 1–54.

[Si] L.Silberman, Addendum by L. Silberman To "Random walk in random group" by M. Gromov, *GAFA* 13 (2003), no. 1, 147-177.

[vDW] van den Dries, L.; Wilkie, A. J. Gromov's theorem on groups of polynomial growth and elementary logic. *J. Algebra* 89 (1984), no. 2, 349–374.

[Z1] R.J. Zimmer, Volume preserving actions of lattices in semisimple groups on compact manifolds, *Pub. Math. IHES*, 59 (1984) 5-33.

[Z2] R.J.Zimmer, Lattices in semisimple groups and distal geometric structures, *Invent. Math.* 80 (1985) 123-137.

[Z3] R.J.Zimmer, Lattices in semisimple groups and invariant geometric structures on compact manifolds, *Discrete Groups in Geometry and Analysis: Papers in Honor of G.D.Mostow*, ed. Roger Howe, Birkhäuser, Boston, 1987.

[Zk] A.Zuk, Property (T) and Kazhdan constants for discrete groups, *GAFA* 13 (2003) no. 3 643-670.

David Fisher
Department of Mathematics and Computer Science
Lehman College - CUNY
250 Bedford Park Boulevard W
Bronx, NY 10468
dfisher@lehman.cuny.edu

Gregory Margulis
Department of Mathematics
Yale University
P.O. Box 208283
New Haven, CT 06520
margulis@math.yale.edu